

DTIC COPY

Naval Ocean Research and Development Activity

March 1990

Report 238



A Global Approach to Image Texture Analysis

AD-A223 532

DTIC
ELECTE
JUN 28 1990
S E D
C E

R. J. Holyer
Ocean Sensing and Prediction Division
Ocean Science Directorate

Foreword

Remote sensing by satellite-borne sensors is one of the most promising methods for obtaining Arctic ice information from this remote and inaccessible part of the earth. However, extracting ice type, concentration, and thickness, as well as measuring lead statistics from satellite imagery, is difficult. The complications of Arctic remote sensing are not yet resolved to the extent necessary to provide all of the products required to support naval operations in the Arctic. This report presents a new approach to image texture analysis that may have application to future sea-ice analysis. The technique is applied to passive microwave imagery of sea ice, where ice classification capability is demonstrated. This new technique could also be applied to synthetic aperture radar (SAR) imagery, where it may provide ice-type information from the SAR to be flown on the European satellite ERS-1 beginning in 1991.

W B Moseley

W. B. Moseley
Technical Director

J B Tupaz

J. B. Tupaz, Captain, USN
Commanding Officer

Executive Summary

A new approach to image texture analysis is developed. The approach is based on linear unmixing of texture measures calculated over an entire image (called a global approach), as opposed to most present texture analysis techniques that compute texture over small neighborhoods (called a local approach). The new global paradigm is appropriate for images where spatial scales of the texture variability are large with respect to the pixel spacing, thereby making the local approach ineffective. Airborne passive microwave imagery of Arctic sea ice contain textures that vary with ice-type. These ice textures are of the type best treated by the global approach. Sea-ice imagery are used as test data to evaluate the global techniques that are developed. Pure, single ice-type images; synthetic mixtures formed by mosaicking pure ice-type subimages in known proportions; and naturally occurring mixture images are analyzed in the course of the study. Proportions of first-year, second-year, and multiyear ice within mixture images are retrieved with root-mean-square accuracies as low as 0.04 by the new method. This accuracy is adequate to be useful in many Arctic studies, but more important, the global technique seems promising for many other remote sensing and general image processing applications. Research areas that are required to advance the global method are enumerated. The most important advancement in support of the global method would be the development of new image texture measures that exhibit linear properties under mixing operations.



Accession For	
NTIS GRA&I	<input checked="" type="checkbox"/>
DTIC TAB	<input checked="" type="checkbox"/>
Unannounced	<input type="checkbox"/>
Justification	
By	
Distribution/	
Availability Codes	
Dist	Avail and/or Special
A-1	

Acknowledgments

The Naval Ocean Research and Development Activity (NORDA) partially supported this work under a long-term training grant. Additional support has been provided by the Office of Naval Technology, Program Element 62435N.

Dr. Duane Eppler of the NORDA Polar Oceanography Branch supplied the KRMS data used in this analysis. He and Mr. L. Dennis Farmer performed the tedious visual interpretation of these data and coincident aerial photography to provide ice-type ground truth for this data set. Dr. Eppler also provided several valuable discussions during the course of this work and offered helpful suggestions after review of the first draft of the manuscript.

Thanks are extended to Dr. William Full of Wichita State University, who provided the CABFAC and QMODEL computer codes.

This work served as the author's Ph.D. dissertation in the Department of Geological Sciences at the University of South Carolina. As a consequence, I wish to thank Dr. Ian Lerche, my major professor, and the other members of my graduate committee: Dr. Robert Ehrlich, Dr. Douglas F. Williams, Dr. Alan E. M. Naion, and Dr. Sterling J. Crabtree.

Contents

Synopsis	1
Summary	1
Conclusions and Recommendations	2
The Global Texture Analysis Paradigm	2
Sea-ice Science	3
Appendix: A Global Approach to Image Texture Analysis	5

A Global Approach to Image Texture Analysis

Synopsis

Texture is an important factor in the analysis of many types of imagery. One category of imagery of special interest to the Navy is remotely sensed imagery of Arctic sea ice. Various types of sea ice are characterized by different textures in ice imagery, whether the images are from a camera, an infrared scanner, a microwave scanner, or imaging radar. Image texture analysis, both visual and computerized is therefore a major tool used by the Arctic science community.

The European Space Agency (ESA) is launching the ERS-1 satellite in 1991. The satellite will carry a synthetic aperture radar (SAR) into a polar orbit. The National Aeronautics and Space Administration (NASA) is building a receiving and processing facility at the University of Alaska in Fairbanks for ERS-1 SAR data from the Arctic region. The Navy plans to use this SAR data at the Naval Polar Ocean Center in Suitland, Maryland, as a supplemental source of ice information to improve its operational Arctic products. One of the key technologies required to insure effective utilization of SAR data is the ability to classify ice types based on texture within the SAR images. Therefore, this study in the appendix is timely in terms of potential Navy application.

Other groups, principally the Environmental Research Institute of Michigan and the Canadian Center for Remote Sensing, are also conducting ice classification studies on the basis of image texture. An important follow-on to this study will be the comparison of the texture analysis techniques developed here with those developed by these other researchers.

Summary

The appendix describes a new approach to image texture analysis and presents the application of that technique to a set of sample imagery. The image data set is airborne passive microwave images of Arctic sea ice. The products of this study are, therefore, twofold. First, a new type of image texture analysis is developed that may have broad application to many branches of science, such as remote sensing, geology, medical

imaging, robotics, and others. Second, in the process of evaluating the texture analysis with the sea-ice data set, new insight is gained concerning the relationships between ice image texture and the age and morphology of sea ice.

Texture is normally considered to be a local property within an image. In other words, texture at a given location in an image can be measured by examining a small area immediately surrounding the point in question. Following this logic, most texture analysis algorithms are based on small, local (neighborhood) calculations. There are, however, a class of textures where the spatial scales of the texture are large with respect to sampling interval, or pixel size within the image. Textures are also found that require calculation of texture statistics over large areas of the image in order to accomplish reliable discrimination of texture types. The ice imagery analyzed in the appendix is representative of this class of textures, which cannot be adequately discriminated using small neighborhood (e.g., 3×3 pixel) operators. However, if the neighborhood size is increased to achieve more reliable estimates of low-frequency or spatially variable textures, then the probability increases that the neighborhood will encompass more than a single texture type, thereby confounding the calculated values of texture parameters. The new paradigm for texture analysis introduced here is termed a global, as opposed to a local, analysis. Since large areas are required to obtain stable statistics for the class of textures under consideration, the neighborhood is simply expanded to include the entire image. Texture measures calculated from the entire image are then linearly unmixed to infer the textural composition of the image, which necessarily leads to the observed global texture measure values. Under the global approach, one does not know where in the image a particular type of texture occurs. In essence then, the global approach is a means of trading positional certainty for more accurate overall estimates of composition. In many applications accurate image composition information is more important than precise positional information, so that this trade-off is advantageous.

The appendix begins by describing the scanning system that generated the sea-ice imagery used as test

data for texture analysis. The geometric and radiometric characteristics of that sensor that would impact texture in the recorded imagery are especially highlighted in that section. Analysis procedures required to avoid scanner artifacts in the texture data are established.

The appendix then reviews image texture analysis, contrasts the local and global approaches, and introduces the concept of a "feature space" (a common term in pattern recognition). Six mathematical properties of texture measures, which are desirable in this study, are described and 25 candidate texture measures are defined. The desired characteristics serve as the yardstick for evaluating the candidate texture measures.

The first image data set is then presented. This data set consists of 64×64 -pixel image fragments containing only a single ice type, which has been determined by prior expert interpretation of the images. Values of the candidate texture measures are calculated for this image set, and the resulting numerical values are analyzed to determine to what degree the candidate texture measures match the desired characteristics established in the previous section. Of the 25 candidate texture variables evaluated, only 12 are shown to be suitable for use in global texture analysis, and then only if the mean brightness values are identical for each texture class. For textures with varying means, only six texture variables are found to be useful.

The unmixing of globally calculated texture data to infer textural composition of the image utilizes the CABFAC and QMODEL computer codes, which require that the data matrix have constant row sums. Image texture data are not intrinsically constant-sum. This problem is addressed and a data matrix transformation, which will put the data into a constant-sum form in a manner that reduces the information loss occurring as a result of the constant-sum formation process, is defined.

The mathematical model for linear unmixing is described, and the appropriateness of that model is evaluated using texture measures calculated for the 64×64 -pixel, pure ice-type samples. Once the validity of the linear unmixing model has been established, the pure samples are "unmixed," even though they are not mixtures. The compositions of these image samples should be 100% for the ice class to which they belong and 0% for all other ice classes. Comparison of the actual compositions calculated by unmixing with these known compositions provides the first quantitative estimate of the ice-type discrimination power of global texture analysis.

The next step is the analysis of images that contain more than one ice type. The mixture images analyzed here are synthetic mixtures created by inosaicking the single ice-type images in various proportions. The unmixing process is conducted for this synthetic data set. Various techniques are applied to select end

members for unmixing. The end-member selection techniques are described and compared. Accuracies are calculated for image proportions resulting from the various techniques.

Natural mixtures are examined by using the synthetic mixture data set to investigate questions pertinent to natural mixtures. These questions arise because natural textures have shifts in mean image intensity levels for each texture class. This shift in means was artificially removed in the first analysis of synthetic mixtures. Because of this mean shift, only six texture measures are suitable for natural mixture images. The synthetic mixtures are recreated, this time without adjusting the mean intensity to a uniform value. The analysis shows that the reduced number of texture variables and the differences in the mean levels between texture classes do not significantly degrade unmixing accuracy.

The final step is the analysis of full-frame images containing natural mixtures of ice types. The analysis of natural mixtures in the present study is limited because of lack of appropriate images in the data set. The work done here with natural, full-frame mixture images indicates that root-mean-square, ice-type proportion errors on the order of 0.05 may be possible for simple, natural mixtures.

Conclusions and Recommendations

Conclusions resulting from this study are divided into two categories. The first deals with the global texture analysis paradigm, which was the major thrust of this work. These conclusions cover matters of broad applicability to the field of image processing. The second category deals with information learned about sea ice as a result of using imagery of that type as test data.

The Global Texture Analysis Paradigm

- The global approach to texture analysis was developed and applied to a test data set with sufficiently promising results that it can be concluded that the approach may have merit for many image texture problems.

- Both mandatory and desirable characteristics of texture measures were enumerated. Of 25 candidate texture measures evaluated, none met all of the desirable and mandatory characteristics, and only two—local homogeneity and Laws' energy masks—possessed the mandatory characteristics, which included correlation with ice type and linear behavior under formation of mixtures. Here, of course, the key item is the linear mixing property. Investigation of additional texture measures that exhibit this characteristic would support further exploitation of the global texture analysis approach.

- A procedure was developed and applied here to deal with the nonconstant-sum nature of image texture

variables. Since this work was performed, a variable-sum version of CABFAC and QMODEL has become available. The analysis described should be repeated using these new programs. If the SIFT used here resulted in significant loss of ice-type information, the repeated analysis with the new code would give better results, indicating even more promise for the global techniques.

- The global approach should be applied to other texture data sets to see if it does indeed have broad applicability.

- Further study should be initiated to deal with the problem of complexity occurring in natural images.

- Failure to solve the complexity problem does not eliminate the global method from practical application. There may be certain applications, industrial inspection and control, for example, where the scenes are always simple mixtures and this technique is applicable without presimplification.

- Global unmixing of image texture is inexorably linked to a trade-off. Proportion accuracy can be exchanged for positional uncertainty. If the global area is large, then proportions can be extracted quite accurately, but one does not know where within the global area the particular texture in question is located. Conversely, as the global area is made smaller to pinpoint the location of a feature, the proportion filled by that texture feature can be measured with less certainty. This situation suggests that a pyramid approach could be advantageously applied. Large areas could be analyzed for maximum accuracy followed by analysis of progressively smaller areas in selected parts of the image where it is judged to be advantageous based on criteria that would be unique to a given application.

Sea-ice Science

- Proportions of first-year, second-year, and multiyear ice types in samples extracted from KRMS images can be calculated with useful accuracy using linear unmixing of globally determined texture measures.

- Dr. Duane Eppler of the Naval Ocean Research and Development Activity (1984) and others have classified ice types in the same KRMS data set. Their

classification work involved thresholding on image brightness. They concluded that brightness temperature thresholding alone was not adequate for classification of ice types. The recommendation in that report was that image texture be included, and provided the stimulus for the present study. We demonstrated here that texture can lead to calculation of accurate (accuracy depending on image size) ice-type proportions. Furthermore, the accuracy has been shown to improve by an approximate factor of two when mean brightness levels are added to texture data. Unfortunately, Eppler's previous work (1984) did not give quantitative estimates of image composition accuracies resulting from the thresholding method, so direct comparison could not be made with the present results.

- The proportion errors as a function of image size, which are derived in the appendix, indicate what order of spatial scales are required to adequately characterize ice types. The 312×512 pixel images that resulted in proportion errors of approximately 0.05 represent 6.4 km² of ice surface.

- The most critical hurdle to practical application of the global technique to ice analysis is believed to be the development of a preprocessing step that will identify and simplify areas of complex textural structure.

- Assuming the above can be accomplished, this texture analysis scheme should be applied to another KRMS data set, where adequate representation of the various ice types will permit complete unmixing using only natural images.

- The real future of ice-texture analysis lies with SAR rather than with passive microwave sensors. The KRMS is a research tool, but operationally the SAR is the sensor of the future. The ERS-1 satellite will have a SAR in space in 1991 and for the foreseeable future thereafter NASA is establishing a receiving facility in Fairbanks, Alaska, to receive that data over a large portion of the Arctic. The Navy Polar Oceanography Center in Suitland, Maryland, is planning to use this SAR data to generate its operational ice products for the U.S. Navy. Clearly, if the texture analysis techniques developed here could be extended to SAR imagery, then the potential payoff would be large in terms of practical application.

Appendix

A Global Approach to Image Texture Analysis

Ronald J. Holyer

I.	Introduction	1
II.	K_α -Band Radiometric Mapping System	7
III.	Image Texture Analysis	18
IV.	The Texture Analysis Data Set	62
V.	The Constant-sum Constraint	93
VI.	Unmixing of Pure Samples	111
VII.	Unmixing of Synthetic Mixtures	139
VIII.	Unmixing with Variable Means	170
IX.	Natural Mixtures	190
X.	Conclusions and Recommendations	208
	Appendix A: The Univariate Maximum Bayesian Classifier	219
	Appendix B: Testing Data Set Normality: The Q-Q Plot	223
	Appendix C: Data Set A	225
	Appendix D: Data Set B	233
	Appendix E: Data Set C	236
	Appendix F: Data Set D	239
	Appendix G: Data Set E	243
	Appendix H: Data Set F	245
	Appendix I: Data Set G	247
	Appendix J: Data Set H	249
	Appendix K: Data Set I	251
	Appendix L: Data Set J	255
	Appendix M: Data Set K	257
	Appendix N: Data Set L	259
	Appendix O: Data Set M	261
	Appendix P: Data Set N	263
	Appendix Q: Data Set O	264

A Global Approach to Image Texture Analysis

by

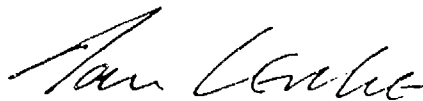
Ronald J. Holyer

Bachelor of Science
Augustana College, 1964

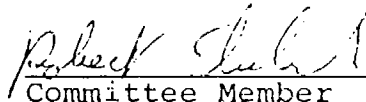
Master of Science
South Dakota School of Mines and Technology, 1966

Submitted in Partial Fulfillment of the Requirements
for the Degree of Doctor of Philosophy in the
Department of Geological Sciences
University of South Carolina

1989



Chairman,
Examining Committee
Major Professor



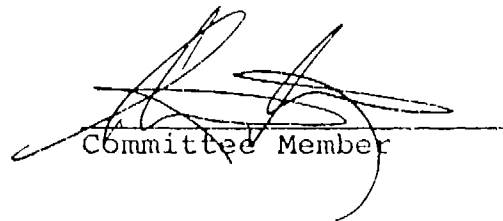
Committee Member



Committee Member



Committee Member



Committee Member

Dean of the Graduate School

Acknowledgements

This study has been supervised by Dr. Ian Lerche, whose guidance has been greatly appreciated. My situation required that much of the work be done while not in residence at the University. Dr. Lerche's willingness to advise me under these inconvenient circumstances was crucial to completion of this study.

Thanks are also due the other members of the committee: Dr. Robert Ehrlich, Dr. Douglas F. Williams, Dr. Alan E. M. Nairn, and Dr. Sterling J. Crabtree.

The Naval Ocean Research and Development Activity (NORDA) partially supported this work under a long-term training grant. Additional support has been provided by the Office of Naval Technology, Program Element 62435N.

Dr. Duane Eppler of the NORDA Polar Oceanography Branch supplied the KRMS data used in this analysis. He and Mr. L. Dennis Farmer performed the tedious visual interpretation of this data and coincident aerial photography, in order to provide ice-type ground truth for this data set. Dr. Eppler also provided several valuable discussions during the course of this work and offered helpful suggestions after review of the first draft of the manuscript.

Thanks are extended to Dr. William Full of Wichita State University who provided the CABFAC and QMODEL computer codes.

I would also like to thank my wife Dorothy, whose support made this work possible.

ABSTRACT

A new approach to image texture analysis is developed. The approach is based on linear unmixing of texture measures calculated over an entire image (called a global approach), as opposed to most present texture analysis techniques that compute texture over small neighborhoods (called a local approach). The new global paradigm is appropriate for images where spatial scales of the texture variability are large with respect to the pixel spacing, thereby making the local approach ineffective. Airborne passive microwave imagery of Arctic sea ice contain textures that vary with ice-type. These ice textures are of the type best treated by the global approach. Sea-ice imagery are used as test data to evaluate the global techniques that are developed. Pure, single ice-type images; synthetic mixtures formed by mosaicking pure ice-type subimages in known proportions; and naturally occurring mixture images are analyzed in the course of the study. Proportions of first-year, second-year, and multiyear ice within mixture images are retrieved with root-mean-square accuracies as low as 0.04 by the new method. This accuracy is adequate to be useful in many Arctic studies, but more important, the global technique seems promising for many other remote sensing and general image processing applications. Research areas that are required to advance the global method are enumerated. The most important advancement in support of the global method would be the development new image texture measures that exhibit linear properties under mixing operations.

TABLE OF CONTENTS

SECTION	Page
ACKNOWLEDGEMENTS	i
ABSTRACT	ii
TABLE OF CONTENTS	iii
LIST OF FIGURES	vii
LIST OF TABLES	xii
I. INTRODUCTION	1
II. K _a -BAND RADIOMETRIC MAPPING SYSTEM	7
A. SPATIAL SAMPLING	7
B. CALIBRATION	11
C. ANTENNA BALANCE	12
D. EFFECT OF OBSERVATION ANGLE ON BRIGHTNESS TEMPERATURE	13
E. KRMS SUMMARY	16
III. IMAGE TEXTURE ANALYSIS	18
A. LOCAL AND GLOBAL APPROACHES TO IMAGE TEXTURE	20
B. DESIRABLE PROPERTIES OF TEXTURE MEASURES	24
1. Correlation with Ice Type	24
2. Multivariate Normality	25
3. Linear Mixing	27
4. Constant Row Sums	28
5. Invariance Under Linear Transformation	29
6. Computational Simplicity	31
C. CANDIDATE TEXTURE MEASURES	31
1. First-Order Statistics	33
2. Gray Level Co-occurrence Matrix	35
3. Laws Masks	45
4. Gradient Analysis	51
5. Run Length	56
6. Summary of Candidate Texture Measures	60
IV. THE TEXTURE ANALYSIS DATA SET	62
A. KRMS SAMPLE IMAGES EXTRACTED AS DATA SET A	62

B. NORMALITY	74
C. CORRELATION WITH ICE TYPE	76
D. LINEAR MIXING	83
E. INVARIANCE UNDER LINEAR TRANSFORMATION	89
F. SUMMARY OF THE TEXTURE ANALYSIS DATA SET	91
V. THE CONSTANT-SUM CONSTRAINT	93
A. THE SCALE INVERT THEN FORCE TRANSFORMATION	98
1. The Concept	98
2. Extension to Higher Dimensions	107
3. Data Set B	107
VI. UNMIXING OF PURE SAMPLES	111
A. THE LINEAR MIXING MODEL	111
B. TRUE DIMENSIONALITY OF DATA SET A	113
1. Principal Components	114
2. Q-mode Factor Analysis	118
C. THE VARIMAX TRANSFORMATION	123
D. END MEMBER PROPORTIONS IN DATA SET B	124
1. Unmixing Data Set B	127
2. Noise Reduction	129
3. Polytope Shape Improvement	130
4. Combined Noise Reduction and Polytope Improvement	134
5. Conclusions from Unmixing Pure Samples	136
VII. UNMIXING OF SYNTHETIC MIXTURES	139
A. SYNTHETIC MIXTURE DATA SET	139
B. UNMIXING WITH CLASS MEANS AS END MEMBERS	147
1. Texture Measures Only	147
2. Texture Plus MEAN	150
C. UNMIXING WITH EXTREME SAMPLES AS END MEMBERS	151
1. Texture Measures Only	154
2. Texture Plus MEAN	157
D. UNMIXING WITH DENEG END MEMBERS	157
1. Texture Measures Only	160
2. Texture Plus MEAN	165

E. SUMMARY OF ANALYSIS OF SYNTHETIC MIXTURES	168
VIII. UNMIXING WITH VARIABLE MEANS	170
A. REDUCTION TO SIX VARIABLES	170
1. Texture Measures Only	171
2. Texture Plus MEAN	174
B. MIXTURES WITH MEAN SHIFTS	177
1. Texture Measures Only	178
2. Texture Plus MEAN	183
3. Unmixing with All Variables of Data Set I	186
IX. NATURAL MIXTURES	190
A. SIMPLE MIXTURES	190
1. Building the Polytope from Large Synthetic Mixtures	190
2. Two End Member Analysis of the Natural Mixtures	198
3. Unmixing Accuracy as a Function of Image Size	201
B. COMPLEX MIXTURES	203
X. CONCLUSIONS AND RECOMMENDATIONS	208
A. THE GLOBAL TEXTURE PARADIGM	208
B. SEA ICE SCIENCE	209
XI. REFERENCES	212
APPENDIX	
A. THE UNIVARIATE MAXIMUM BAYESIAN CLASSIFIER	219
B. TESTING DATA SET NORMALITY: THE Q-Q PLOT	223
C. DATA SET A	225
D. DATA SET B	233
E. DATA SET C	236
G. DATA SET E	243
H. DATA SET F	245
I. DATA SET G	247
J. DATA SET H	249
K. DATA SET I	251

L. DATA SET J	255
M. DATA SET K	257
N. DATA SET L	259
O. DATA SET M	261
P. DATA SET N	263
Q. DATA SET O	264

LIST OF FIGURES

FIGURE	Page
I-1. Typical 33.6 GHz image of Arctic sea ice.	2
II-1. KRMS scanning geometry.	9
II-2. Relative size and shape of the KRMS IFOV.	9
II-3. Comparison of Fresnel model of vertically polarized microwave brightness temperatures with data from KRMS imagery.	15
III-1. A simple structured texture.	19
III-2. Hypothetical data set in a two-dimensional feature space.	21
III-3. Representative samples of KRMS imagery of first-year, second-year, and multiyear ice.	22
III-4. Two binary images and their corresponding GLC matrices.	37
III-5. Two visually distinct textures which have identical second-order probabilities.	39
III-6. The GLC matrix-based inertia texture measure as a function of displacement vector length for three typical 128 × 128 pixel samples of FY, SY, and MY ice.	42
III-7. The GLC matrix-based correlation texture measure as a function of displacement vector length for three typical 128 × 128 pixel samples of FY, SY, and MY ice.	42
III-8. Center-weighted vector masks.	47
III-9. Normalized energy density spectra of the seventh-order masks shown in Figure III-8.	50
III-10. Center-weighted vector masks.	52
III-11. Normalized energy density spectra of the 21st order masks shown in Figure III-10.	53
III-12. Illustration of image energy spectrum partitioning by spatial convolution of a 128 × 128 pixel KRMS sample of multiyear ice with the masks given in Figure III-10.	54
III-13. Representative KRMS images of first-year and multiyear ice with their corresponding edge magnitude images.	57
III-14. Representative KRMS images of first-year and multiyear ice with their corresponding binary images formed by thresholding the image at its mean intensity value.	59

IV-1. Full 512×512 pixel image showing locations where 64 \times 64 pixel subareas were extracted as samples MY1 through MY7 and FY5 of data set A.	63
IV-2. Full 512×512 pixel image showing locations where 64 \times 64 pixel subareas were extracted as samples MY8 and MY9 of data set A.	64
IV-3. Full 512×512 pixel image showing locations where 64 \times 64 pixel subareas were extracted as samples MY10 through MY12 and FY6 through FY7 of data set A.	65
IV-4. Full 512×512 pixel image showing locations where 64 \times 64 pixel subareas were extracted as samples SY1 through SY12 of data set A.	66
IV-5. Full 512×512 pixel image showing locations where 64 \times 64 pixel subareas were extracted as samples FY1 through FY4 of data set A.	67
IV-6. Full 512×512 pixel image showing locations where 64 \times 64 pixel subareas were extracted as samples MY13 through MY18 of data set A.	68
IV-7. Full 512×512 pixel image showing locations where 64 \times 64 pixel subareas were extracted as samples SY13 and SY14 of data set A.	69
IV-8. Full 512×512 pixel image showing locations where 64 \times 64 pixel subareas were extracted as samples SY15 through SY18 of data set A.	70
IV-9. Full 512×512 pixel image showing locations where 64 \times 64 pixel subareas were extracted as samples SY19 through SY22 of data set A.	71
IV-10. Full 512×512 pixel image showing locations where 64 \times 64 pixel subareas were extracted as samples SY23 through SY25 of data set A.	72
IV-11. Histogram of linear correlation coefficient values for Q-Q plots of data from data set A.	75
IV-12. A FY/MY mixture image.	84
IV-13. A FY/MY mixture image with the mean brightness of each half offset to a value of 128.	87
V-1. One-dimensional linear mixing of a variable row sum data set, and a constant row sum data set.	94
V-2. Plot of the hypothetical data set listed in Table V-1.	102
V-3. Illustration of rotation and constant-sum forcing of the data shown in Figure V-2.	102

V-4. Graphical representation of the scale, invert, and force transformation (SIFT) of the data shown in Figure V-2. . . .	104
VI-1. Coefficients of determination for the 12 variables of data set B as a function of number of end-members. . . .	122
VI-2. Plot of data set B in the reduced feature space. . . .	125
VI-3. Plot of data set B in the reduced feature space with variables LM4 and LM5 excluded from the analysis. . . .	131
VI-4. Plot of data set B in the reduced feature space with MEAN included along with the 12 original texture variables. . . .	133
VI-5. Plot of data set C in the reduced feature space. . . .	135
VI-6. Coefficients of determination for the 11 variables of data set C as a function of number of end members. . . .	137
VII-1. Synthetic mixture images MX1 - MX6 produced by merging pure image samples from data set A. . . .	140
VII-2. Synthetic mixture images MX7 - MX12 produced by merging pure image samples from data set A. . . .	141
VII-3. Synthetic mixture images MX13 - MX18 produced by merging pure image samples from data set A. . . .	142
VII-4. Synthetic mixture images MX19 - MX24 produced by merging pure image samples from data set A. . . .	143
VII-5. Synthetic mixture images MX25 - MX30 produced by merging pure image samples from data set A. . . .	144
VII-6. Synthetic mixture images MX31 - MX36 produced by merging pure image samples from data set A. . . .	145
VII-7. Synthetic mixture images MX37 - MX42 produced by merging pure image samples from data set A. . . .	146
VII-8. Plot of data set E in the reduced feature space. . . .	148
VII-9. Plot of data set F in the reduced feature space. . . .	152
VII-10. Plot of data set E in the reduced feature space. Solid dots indicate Imbrie end members. . . .	155
VII-11. Plot of data set F in the reduced feature space. Solid dots indicate Imbrie end members. . . .	158
VII-12. Error in calculated sample proportions as a function of DENEG cut-off value for data set F. . . .	161
VII-13. Plot of data set E in the reduced feature space. Solid dots indicate DENEG end members. $t_2 = -0.25$	162
VII-14. Plot of data set E in the reduced feature space. Solid dots indicate DENEG end members. $t_2 = -0.4$	163

VII-15. Error in calculated sample proportions as a function of DENEG cut-off value for data set E.	166
VII-16. Plot of data set F in the reduced feature space. Solid dots indicate DENEG end members. $t_2 = -0.325$	167
VIII-1. Plot of data set G in the reduced feature space.	172
VIII-2. Plot of data set G in the reduced feature space with various superimposed mixing polytopes.	173
VIII-3. Plot of data set H in the reduced feature space.	175
VIII-4. Plot of data set H in the reduced feature space with various superimposed mixing polytopes.	176
VIII-5. Typical synthetic mixture image (NX2) without mean intensity level adjusted prior to mosaicking.	179
VIII-6. Plot of data set J in the reduced feature space.	180
VIII-7. Plot of data set J in the reduced feature space with various superimposed mixing polytopes.	182
VIII-8. Plot of data set K in the reduced feature space.	184
VIII-9. Plot of data set K in the reduced feature space with various superimposed mixing polytopes.	185
VIII-10. Plot of data set L in the reduced feature space.	187
VIII-11. Plot of data set L in the reduced feature space with various superimposed mixing polytopes.	188
IX-1. KRMS image BX1 containing a natural mixture of FY (90%) and MY (10%) ice.	191
IX-2. KRMS image BX2 containing a natural mixture of FY (40%) and MY (60%) ice.	192
IX-3. KRMS image BX3 containing a natural mixture of FY (69%) and MY (31%) ice.	193
IX-4. KRMS image BX4 containing a natural mixture of FY (60%) and MY (40%) ice.	194
IX-5. Synthetic KRMS mixture image consisting entirely of MY ice.	195
IX-6. Plots of data set N in the reduced feature space. Ploytope is constructed by connecting large synthetic mixtures of a single ice type.	197
IX-7. Plots of data set O in the reduced feature space. Ploytope is constructed by connecting large synthetic mixtures of a single ice type.	200

IX-8. Calculated errors in unmixing proportions as a function of image size. The dashed curve is a S^{-1} reference line.	202
IX-9. Complex KRMS natural mixture of FY and MY ice. Boxes denote a solid MY flow and MY rubble.	204
IX-10. Natural mixture image of Figure IX-9 with complex rubble areas covered with a uniform gray level.	206
A-1. Distributions of the Laws' Mask 2 texture measure values from data set A for first-year, second-year, and multiyear ice types.	222
B-1. Q-Q plots for the Laws' Mask 2 texture measure calculated from data set A.	224

LIST OF TABLES

TABLE	Page
II-1. KRMS TECHNICAL CHARACTERISTICS	8
III-1. CANDIDATE IMAGE TEXTURE MEASURES	61
IV-1. DATA SET A STATISTICS	78
IV-2. BIVARIATE ICE TYPE CLASSIFICATION ACCURACIES . .	81
IV-3. TEXTURE MEASURES FROM FIG. IV-12	85
IV-4. TEXTURE MEASURES FROM FIG. IV-13	88
IV-5. EFFECT OF INTENSITY TRANSFORMATION ON IMAGE TEXTURE MEASURES	90
V-1. A HYPOTHETICAL TWO-DIMENSIONAL DATA SET . . .	99
V-2. DATA FROM TABLE V-1 AFTER THE FRF TRANSFORMATION	101
V-3. DATA FROM TABLE V-1 WITH THE SIFT APPLIED . . .	105
V-4. CORRELATION OF TEXTURE VARIABLES WITH VAR	108
VI-1. EIGENVALUES OF C FROM DATA SET B	117
VI-2. EIGENVALUES OF P FROM DATA SET B	121
VII-1. UNMIXING RESULTS FOR DATA SETS E AND F - Class Means from Sets E and F as End Members	149
VII-2. UNMIXING RESULTS FOR DATA SETS E AND F - Imbrie End Members	156
VII-3. UNMIXING RESULTS FOR DATA SETS E AND F - DENEG End Members	164
C-1. VALUES OF TEXTURE MEASURES FOR DATA SET A . . .	226
D-1. VALUES OF TEXTURE MEASURES FOR DATA SET B . . .	234
E-1. DATA SET C	237
F-1. VALUES OF TEXTURE MEASURES FOR DATA SET D . . .	240
G-1. DATA SET E	244
H-1. DATA SET F	246
I-1. DATA SET C	248
J-1. DATA SET H	250
K-1. VALUES OF TEXTURE MEASURES FOR DATA SET I . . .	252
L-1. DATA SET J	256

M-1. DATA SET K	258
N-1. DATA SET L	260
O-1. VALUES OF TEXTURE MEASURES FOR DATA SET M . .	262
P-1. DATA SET N	263
K-1. VALUES OF TEXTURE MEASURES FOR DATA SET I . . .	264

I. INTRODUCTION

This dissertation describes a new approach to image texture analysis and presents the application of that technique to a set of sample imagery. The image data set is airborne passive microwave images of Arctic sea ice, such as shown in Figure I-1. The products of this study are, therefore, twofold. First, a new type of image texture analysis is developed that may have broad application to many branches of science, such as remote sensing, geology, medical imaging, robotics, and others. Second, in the process of evaluating the texture analysis with the sea-ice data set, new insight is gained concerning the relationships between ice image texture and the age and morphology of sea ice.

Texture is normally considered to be a local property within an image. One definition of texture is the "apparent minute pattern of detail in a given area" (Hsu, 1979). The words "minute" and "detail" in this definition imply a local character to image texture. In other words, texture at a given location in an image can be measured by examining a small area immediately surrounding the point in question. Following this logic, most texture analysis algorithms are based on small, local (neighborhood) calculations. There are, however, a class of textures where the spatial scales of the texture are large with respect to sampling interval, or pixel size within the image. Textures are also found which require calculation of texture statistics over large areas of the image in order to accomplish reliable discrimination of texture types. The ice imagery analyzed here is representative of this class of textures that cannot be adequately discriminated using small neighborhood (*e.g.*, 3×3 pixel) operators. However, if the neighborhood size is increased to achieve more reliable estimates of low-frequency or spatially variable textures, then the probability increases that the neighborhood will en-



FIGURE 1-1. Typical 33.6 GHz image of Arctic sea ice.

compass more than a single texture type, thereby confounding the calculated values of texture parameters. The new paradigm for texture analysis introduced here is termed a global, as opposed to a local, analysis. Since large areas are required to obtain stable statistics for the class of textures under consideration, the neighborhood is simply expanded to include the entire image. Texture measures calculated from the entire image are then linearly unmixed to infer the textural composition of the image, which necessarily leads to the observed global texture measure values. Under the global approach, one does not know where in the image a particular type of texture occurs. In essence then, the global approach is a means of trading positional certainty for more accurate overall estimates of composition. In many applications accurate image composition information is more important than precise positional information, so that this trade-off is advantageous.

The following paragraphs present an overview of this dissertation.

Section II describes the scanning system that generated the sea-ice imagery used as test data for texture analysis. The geometric and radiometric characteristics of that sensor that would impact texture in the recorded imagery are especially highlighted in that section. Analysis procedures required to avoid scanner artifacts in the texture data are established.

Section III reviews image texture analysis, contrasts the local and global approaches, and introduces the concept of a "feature space" (a common term in pattern recognition). Six mathematical properties of texture measures, which are desirable in this study, are described and 25 candidate texture measures are defined. The desired characteristics serve as the yardstick for evaluating the candidate texture measures.

Section IV presents the first image data set. This data set consists of 64×64 -pixel image fragments containing only a single ice type, which has been determined by prior expert interpretation of the images. Values of the candidate texture measures are calculated for this image set, and the resulting numerical values are analyzed to determine to what degree the candidate texture measures match the desired characteristics established in the previous section. Of the 25 candidate texture variables evaluated, only 12 are shown to be suitable for use in global texture analysis, and then only if the mean brightness values are identical for each texture class. For textures with varying means, only six texture variables are found to be useful.

The unmixing of globally calculated texture data to infer textural composition of the image utilizes the CABFAC and QMODEL computer codes, which require that the data matrix have constant row sums. Image texture data are not intrinsically constant-sum. Section V addresses this problem. A data matrix transformation, which will put the data into a constant-sum form in a manner that reduces the information loss occurring as a result of the constant-sum formation process, is defined.

Section VI introduces the linear unmixing procedure that is used to infer image composition from global data. The mathematical model for linear unmixing is described and the appropriateness of that model is evaluated using texture measures calculated for the 64×64 -pixel, pure ice-type samples from Section IV. Once the validity of the linear unmixing model has been established, the pure samples are "unmixed," even though they are not mixtures. The compositions of these image samples should be 100% for the ice class to which they belong and 0% for all other ice classes. Comparison of the actual compositions calculated by unmixing with these known compositions provides the first quantitative estimate

of the ice-type discrimination power of global texture analysis.

Section VII presents the first analysis of images that contain more than one ice-type. The mixture images analyzed here are synthetic mixtures created by mosaicking the single ice-type images in various proportions. The unmixing process is conducted for this synthetic data set. Various techniques are applied to select end members for unmixing. The end-member selection techniques are described and compared. Accuracies are calculated for image proportions resulting from the various techniques.

Section VIII looks toward natural mixtures by using the synthetic mixture data set to investigate questions pertinent to natural mixtures. These questions arise because natural textures have shifts in mean image intensity levels for each texture class. This shift in means was artificially removed from the synthetic mixtures analyzed in Section VII. Because of this mean shift, only six texture measures are suitable for natural mixture images. The synthetic mixtures are recreated, this time without adjusting the mean intensity to a uniform value. The analysis shows that the reduced number of texture variables and the differences in mean levels between texture classes do not significantly degrade unmixing accuracy.

Section IX contains the analysis of full frame images containing natural mixtures of ice types. The analysis of natural mixtures in the present study is limited because of lack of appropriate images in the data set. What work is done here with natural, full-frame mixture images indicates that root-mean-square (rms) ice-type proportion errors on the order of 0.05 may be possible for simple, natural mixtures.

The conclusions of this study are stated in Section X. Recommendations

are presented for further study to advance the global texture analysis approach. Recommendations are also presented for future ice work using the global technique. The extension of this approach to synthetic aperture radar (SAR) images of sea ice is seen as especially significant.

II. K_a-BAND RADIOMETRIC MAPPING SYSTEM

The K_a-band Radiometric Mapping System (KRMS) is an airborne microwave imager that operates at a center frequency of 33.6 GHz. Important characteristics of the instrument are given in Table II-1. Those characteristics that pertain to texture analysis of imagery collected by the KRMS are discussed in this section.

A. SPATIAL SAMPLING

The 1° antenna beamwidth results in a nadir instantaneous field of view (IFOV) of 16 ft per 1000 ft of flight altitude. Postacquisition signal processing is applied to the data to achieve a cross-track IFOV of 12 ft per 1000 ft of altitude in the digital imagery (Eppler *et al.*, 1984). All imagery analyzed in the present study was acquired from an aircraft altitude of 5000 ft; the result was a 60-ft cross-track by 80-ft along-track IFOV in the digital data.

The active scan angle of the KRMS is 100° centered on nadir. The resultant cross-track ground coverage is equal to 2.38 times the altitude, or 11,900 ft (approximately 2 nmi) for a 5000-ft altitude. If aircraft altitude is low enough that curvature of the earth can be ignored, then scan geometry for an airborne scanner is shown in Figure II-1, where θ is the scan angle relative to nadir, h is the aircraft altitude, P is the ground point being sensed, x is the ground distance from nadir to P , and d is the distance from the sensor to P . Trigonometry leads to an expression for x as a function of θ and h .

$$x = h \tan \theta. \quad (II - 1)$$

If KRMS data were digitized using a constant time interval between sam-

TABLE II-1. KRMS TECHNICAL CHARACTERISTICS
taken from Eppler *et al.* (1984)

ANTENNAS		
	number	3
	diameter	24 inches
	polarization	vertical
	beamwidth	1.0°
	isolation	40 dB (minimum)
SCANNER		
	maximum scan rate	25 scans/second
	minimum scan rate	7.5 scans/second
	scan angle	60° from nadir
	active scan angle	50° from nadir
	antenna position accuracy	2.5 minutes if arc
STABILIZATION		
	method	cross-track roll gyro
	accuracy	less than 0.25°
RF AMPLIFIER		
	type	superheterodyne (DSB)
	noise	less than 5.0 dB
	bandwidth	1.3 GHz
	gain	greater than 60 dB
	loss	1.2 dB (maximum)
RADIOMETER		
	type	pulse stabilized ^d , total power
	pulse width	4.0 ms
	local oscillator frequency	33.6 GHz
	IF bandwidth	greater than 500 MHz
	video bandwidth	1.7 kHz (maximum)
	video gain	72 dB (nominal)
	minimum detectable signal	0.05 K/second
	sensitivity	50 mV/K (nominal)
	dynamic range	370 K

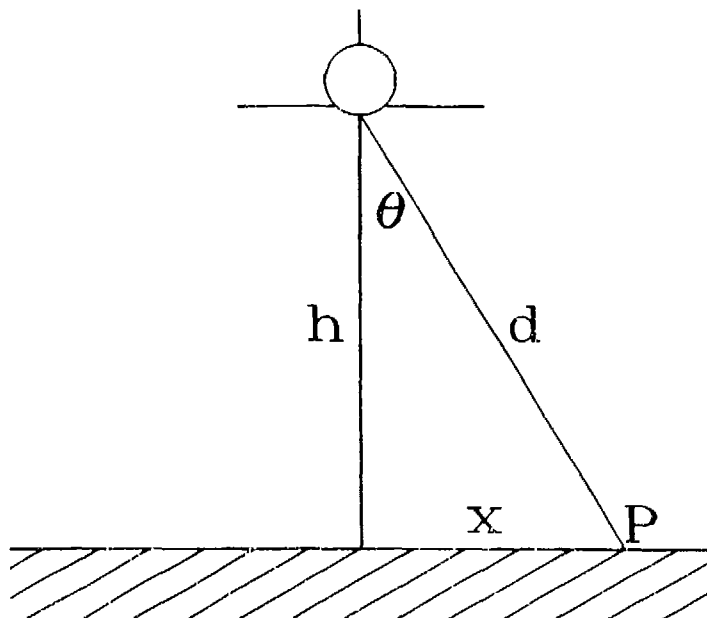


FIGURE II-1. KRMS scanning geometry.

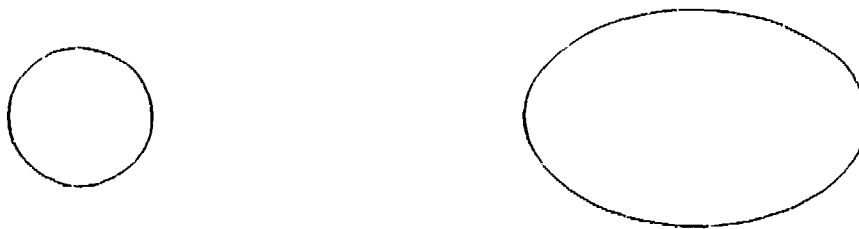


FIGURE II-2. Comparison of the relative size and shape of the KRMS IFOV at nadir (left) and at the scan edge (right).

ples (which is equivalent to a constant angular interval), the sample interval in terms of ground distance would not be uniform across the scan line, but would vary as $\frac{dx}{d\theta}$. Differentiating Eq. (II-1) with respect to θ yields

$$\frac{dx}{d\theta} = h \sec^2 \theta. \quad (II - 2)$$

Substituting θ values of 0° and 50° into Eq. (II-2) shows that the value of $\frac{dx}{d\theta}$ varies by a factor of 2.42 from nadir to the end of the scan line. This amount of variability in sample spacing on the ground would certainly be intolerable in texture analysis where spatial relationships of image intensity are crucial. Fortunately, the analog-to-digital (A/D) conversion software for the KRMS system incorporates a nonuniform sampling in time, appropriately chosen to result in uniform sampling in ground distance.

However, another potential texture analysis problem associated with scanner geometry results from the fact that IFOV is a function of θ . At the edge of the scan the IFOV represents an integration over a much larger ground area than the IFOV at nadir. Both d and θ combine to enlarge the cross-track dimension of the IFOV in proportion to $\frac{dx}{d\theta}$. The along-track dimension of the IFOV is increased in proportion to d which is proportional to $(\frac{dx}{d\theta})^{\frac{1}{2}}$. Figure II-2 compares the KRMS IFOV at nadir and at the edge of the scan. Clearly one would expect significant loss of high-frequency content in the image toward the edge of the scan. The typical KRMS image shown in Figure I-1 exhibits this effect in the form of a fuzzy or defocused appearance at the image edges.

Each KRMS scan line is digitized into 512 samples. Assuming a 5000-ft aircraft altitude where the cross-track coverage would be 11,900 ft, the sampling interval in terms of ground distance between each of the 512 samples in a scan

line would be 23.3 ft. Comparison of this intersample spacing with the IFOV of 60 ft at nadir shows the data to be oversampled by a factor of 2.58 at nadir. The cross-track oversampling factor becomes even larger toward the ends of the scan line where the IFOV increases.

The along-track sampling interval is determined by a combination of aircraft altitude, aircraft ground speed, and antenna rotation rate, all of which can be adjusted in flight by the KRMS operator. Nominal values for these three parameters during the acquisition of the KRMS data analyzed here are 5000-ft altitude, 221-kt ground speed, and 50 ms per scan. Resultant scan line separation on the ground is 18.5 ft. This separation, when compared to the 80-ft along-track IFOV, indicates an oversampling factor of 4.32 in the along-track direction.

The aspect ratio in the KRMS imagery is not unity. Along-track and cross-track sampling intervals of 18.5 ft and 23.3 ft, respectively, result in an aspect ratio of 1.26, elongated in the along-track direction. One could attempt to correct the images for aspect ratio, but such a correction would be subject to uncertainties. Aircraft altitude and speed recorded on the flight log are only estimates that can vary from the true values. Because of these and other uncertainties (such as aircraft crab angle) in platform attitude and motion, no aspect ratio correction to the imagery has been attempted.

B. CALIBRATION

The KRMS data is digitized to 12-bit precision resulting in possible digital values ranging from 0 to 4095. The data have been calibrated by methods described in Eppler *et al.* (1984) so that the equation

$$T_b = (-0.0276 * \text{digital value}) + 248 \quad (II - 3)$$

can be used to convert digital image values to brightness temperature, T_b , in Kelvins (K). In the present study it was convenient to work with 8-bit data, which can assume values ranging from 0 to 255. The 12-bit data were converted to 8 bits by dropping the four least-significant bits. The calibration equation for the resulting 8-bit data is

$$T_b = (-0.4416 * \text{digital value}) + 248. \quad (II - 4)$$

The range represented by the 8 bits is unchanged from 12-bit data (135 to 248 K), but the sensitivity in the 8-bit data is reduced to 0.4416 K/digital value rather than the 0.0276 K/digital value present in the original digital imagery. The reduced sensitivity of the 8-bit data is more than adequate for the texture analysis performed here, since the rms noise level in the KRMS data is approximately 2.5 K. The four least-significant bits that were dropped therefore represent temperature sensitivity that is well below the random noise floor in the data. Note that Eq. (II-4) describes the case where darker shades in the imagery represent higher radiometric temperatures and lighter shades represent lower temperatures.

C. ANTENNA BALANCE

The KRMS scanning system consists of three parabolic antennas mounted 120° apart on a single rotating shaft. Thus, one revolution of the shaft produces three scan lines, one from each of the antennas. Only the downward-looking antenna is active at any given time. The A/D conversion software for the KRMS

data contains provisions to compensate for differences in the response characteristics of the three antennas. This compensation is adequate to insure that scan-line banding is not readily apparent in the images. However, the compensation is not perfect and some sensitive texture measures may be affected by the imbalances within the scan-line triplets.

D. EFFECT OF OBSERVATION ANGLE ON BRIGHTNESS TEMPERATURE

The underlying assumption in this study is that changes in the physical characteristics of the ice pack (e.g., salinity, liquid water content, ice thickness, deformation, etc.) will produce textured distributions of brightness temperatures, T_b , that can be measured from microwave imagery and utilized to identify specific ice types. However, superimposed upon these textured T_b distributions that result from the physical characteristics of the ice, one finds a T_b distribution resulting from the effect of observation angle. These angular effects can confound our attempts to produce accurate measures of that component of image texture related to ice type.

Eppler *et al.* (1984), and references cited therein, discuss the problem of observation angle effects on microwave brightness temperatures of sea ice. The problem can be summarized in simple terms by considering a model treating the ice as a semi-infinite dielectric slab. The dielectric is assumed to be lossy, resulting in a complex dielectric constant, ϵ . If we consider only vertically polarized emissions from the surface (the KRMS senses only the vertically polarized component), the vertical Fresnel coefficient $R_v(\theta)$ for an electromagnetic wave emerging from the dielectric is:

$$R_v(\theta) = \frac{(\epsilon \cos \theta - S)}{(\epsilon \cos \theta + S)}, \quad (II - 5)$$

where

$$S = \epsilon - \sin \theta.$$

In the microwave portion of the spectrum, the intensity of the radiation emitted by a "gray" body is frequency dependent and is proportional to physical temperature T_0 . Since the intensity is proportional to the square of the field strength, we have outside the dielectric (Stogryn, 1970):

$$T_b(\theta) = |R_v(\theta)|^2 T_{sky} + (1 - |R_v(\theta)|^2) T_0. \quad (II - 6)$$

T_{sky} is the radiometric temperature of the sky, which in the microwave portion of the spectrum is typically less than 30 K. If we assume that T_{sky} is 15 K, ϵ for ice is (3.15, 0.011) (Stogryn, 1981) and T_0 is 270 K, then the simple Fresnel model of Eq. (II-6) results in the $T_b(\theta)$ curve shown in Figure II-3. Note that the model predicts an increase in brightness temperature with increasing scan angle. The elevation in T_b relative to the nadir value reaches a maximum of 20 K at an observation angle of approximately 61° . At 30° the model predicts an increase in T_b of approximately 6.4 K. The magnitude of this temperature elevation with angle will vary with ice type because the salinity and, therefore, the complex dielectric constant, varies with the age of the ice.

The KRMS imagery from March 1983 contains a large area of shore-fast, first-year ice that can be assumed to be isothermal. Eppler *et al.* (1984) calculated average T_b as a function of scan angle for this isothermal ice field. Figure II-3 shows those results along with the theoretical curve from Eq. (II-6). Close agreement exists between theory and observations because the chosen

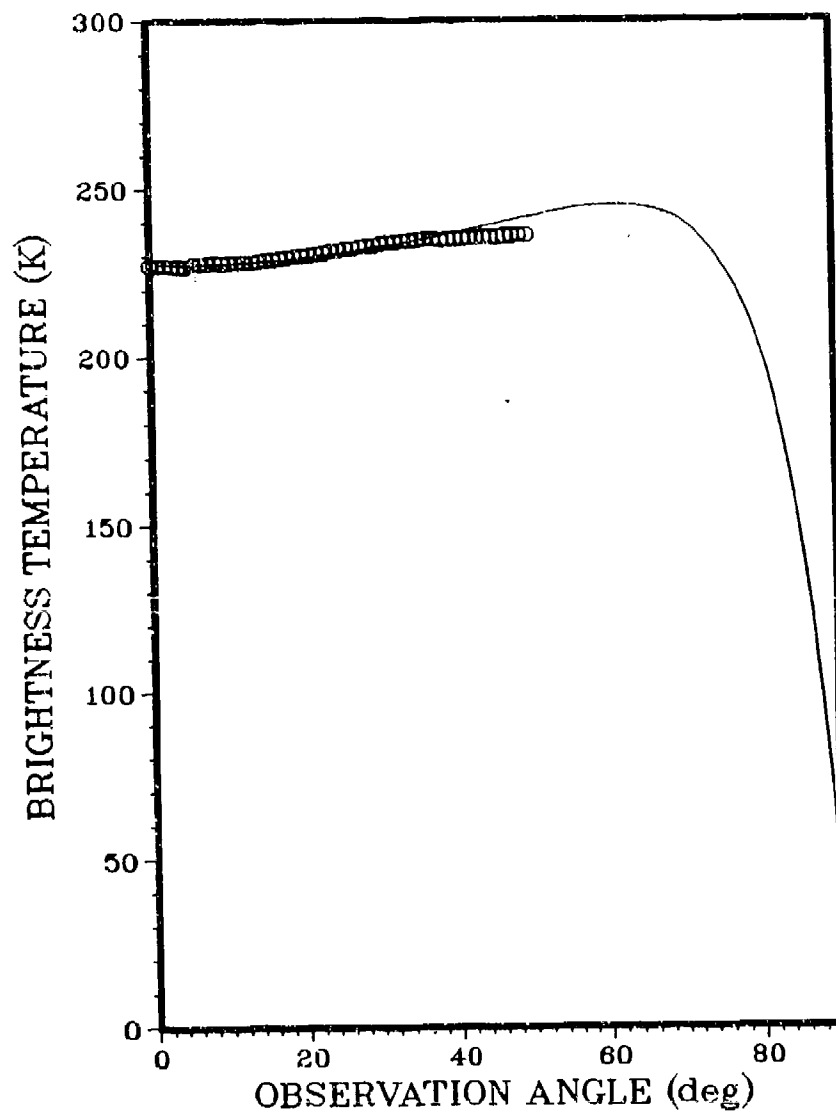


FIGURE II-3. Comparison of Fresnel model of vertically polarized microwave brightness temperature as a function of scan angle (line) with data from KRMS imagery of isothermal ice field (circles).

value for ϵ was apparently appropriate for this first-year ice. The conclusion is that T_b distributions will be a linear combination of texture related to ice type and some limb-warming trend, which is also related to ice type through the complex dielectric constant in Eq. (II-6). The procedure to subtract the limb trend from the KRMS imagery is not straightforward, since the magnitude of the trend is different for different ice types. Therefore, correction for this ramp has not been included in the present study. This correction is suggested as a possible enhancement to future work. Eppler (1987) is studying this problem.

E. KRMS SUMMARY

Characteristics of the KRMS have been discussed here to be certain that the analysis is undertaken with full recognition of any scanner-related factors that could have significant impact on the textures of the images recorded by this instrument. To summarize the factors discussed, the following potential problems and possible solutions are listed.

- The aspect ratio of the images is 1.26, elongated in the along-track direction. No correction for aspect ratio has been performed. The problem will be ignored, but textures will not be isotropic.
- Data are oversampled by a factor that ranges from 2.58 to 6.24, depending on direction within the image and distance from the image edge. This problem will be reduced by eliminating the 100 edge pixels on either side of the image (where the oversampling is most severe) from inclusion in the study. Excluding these edge pixels reduces the range of the oversampling factor to 2.58 to 4.32. It is noted that this oversampling will result in an absence of high-frequency content in the imagery. Texture measures associated with high frequencies are expected to be of no value.

- Each group of three successive scan lines is sensed through a different antenna. Imbalance in the antenna response characteristics may lead to a periodic ($f = 0.33/\text{pixel}$) texture artifact in the along-track direction. Texture measures will be implemented in such a way that the effects of this artifact are minimized.
- The images analyzed here are calibrated in terms of microwave brightness temperature. Calibration is an advantage for texture analysis because we can ignore image normalization by histogram equalization or by other means which are normally required prior to texture analysis (Haralick *et al.*, 1973).
- Brightness temperatures increase toward the edge of the scan line because of the effects of incidence angle. Issues of resolution have led us to discount the 100 edge pixels on either side of the imagery. The elimination of these edge pixels also minimizes the limb warming problem. By ignoring the 100 edge pixels, the remaining image area is observed at an angle of 30° or less where the limb warming is less severe than at the ends of the scan line. It is assumed that this cross-track brightness trend can be ignored in the subsequent texture analysis. This assumption would certainly be valid for texture analysis based on the local neighborhood. However, under the global approach to texture analysis outlined in Section III, the expected cross-track trend will undoubtedly have some effect.

III. IMAGE TEXTURE ANALYSIS

Texture is an important factor in the analysis of many types of images. Despite its importance in image data analysis, a precise mathematical definition of texture does not exist. Texture has been defined subjectively as the "local spatial distribution of tonal values within an image" (Haralick *et al.*, 1973), or the "apparent minute pattern of detail of a given area," (Hsu, 1979). Because of the lack of a mathematical definition of texture, it is difficult to pursue a rigorous, formal approach to texture analysis. Texture discrimination techniques have therefore been described as being, for the most part, *ad hoc* (Haralick, 1979). However, two broad classes of textures are generally recognized. These texture classes are called structured and statistical.

Structured textures are characterized by texture primitives and placement rules. Figure III-1 is an example of of a structured texture. Here the texture primitive is the pattern T_EX, and the placement rule is for a primitive to occur at evenly spaced grid points. Structured textures abound - woven fabric, brick walls, and wire mesh are typical examples. Structured textures are popular subjects for study because the primitives and placement rules can be formalized into grammatical models. Through a small number of rules and symbols the grammar can generate complex textural patterns. The structural approach is also attractive because it brings some mathematical formalism to bear on the texture discrimination problem where none exists naturally. For an overview of the application of grammatical models to structured texture analysis, see Ballard and Brown (1982).

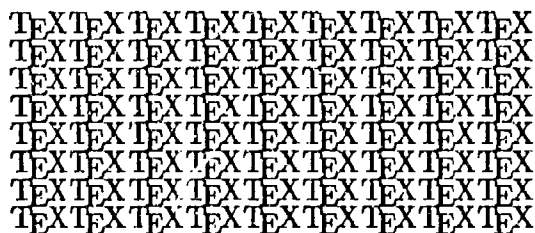


Figure III-1. A simple structured texture constructed using `TeX` as the texture primitive and a rectangular grid as the placement rule.

Textures that cannot be adequately described by primitives and placement rules fall into the statistical class where texture is described by statistical rules governing the spatial distribution and relation of tonal values. In the field of remote sensing nearly all natural textures observed in aerial images of the earth fall into the statistical category. A sample of the KRMS imagery studied in this investigation was shown in Figure I-1. Recall the lack of any structure to the texture contained in the image. Analysis of these data will, therefore, follow a statistical approach. Specifically, statistical pattern recognition is the paradigm that will be used to explore sea ice textures. There are other statistical methods of dealing with texture (Pratt *et al.* 1981), but this study will be confined to the statistical pattern recognition approach, which seems to be particularly appropriate for low-resolution textures such as those seen in aerial images (Weszka *et al.*, 1976).

The basic notion of pattern recognition is the "feature vector." A set of m image measurements, x_i , is considered to be a feature vector, \vec{X} ,

$$\vec{X} = (x_1, x_2, x_3, \dots, x_m)$$

in a "feature space" of m dimensions. For texture analysis the image measurements should, of course, be measures of texture type so that feature vectors

cluster according to the texture from which they were derived. Figure III-2 shows examples of effective and ineffective feature selection. If features can be found such that the feature vectors cluster by texture type in feature space, then the methods of statistical pattern recognition can be brought to bear on the image texture analysis problem.

A. LOCAL AND GLOBAL APPROACHES TO IMAGE TEXTURE

Because texture is usually considered to be the local distribution of tonal values, most image texture studies take a local approach to texture analysis. Under the local approach, one determines the texture class of a given pixel by examining the pixels in a neighborhood surrounding the pixel in question. If one wanted to know the percentage of an image consisting of some texture type, each pixel in the image would be classified based on its neighborhood and then the number of pixels classified as belonging to a given class would be summed. Examples of the local approach are abundant in the literature. For example, Hsu (1979) used both 3×3 and 5×5 pixel neighborhoods for texture-based classification of terrain types in aerial imagery. Holmes *et al.* (1984) have classified SAR imagery of sea ice using a circular neighborhood with a 5 pixel (15 m) radius. Fily and Rothrock (1986) used a 15×15 pixel neighborhood to discriminate between ice and open water in SAR imagery. Laws (1980) segmented images based on texture by convolution with small masks ranging from 3×3 to 7×7 pixels.

The local approach is not optimal for the sea ice texture analysis at hand. Consider the image examples shown in Figure III-3. Each ice type clearly has a different textural appearance. However, if one extracted a 3×3 or 5×5 pixel neighborhood from the second-year or multiyear ice images, then it would

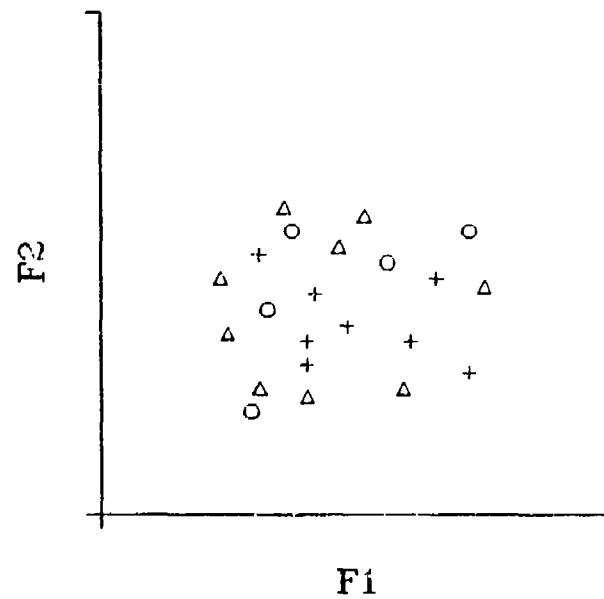
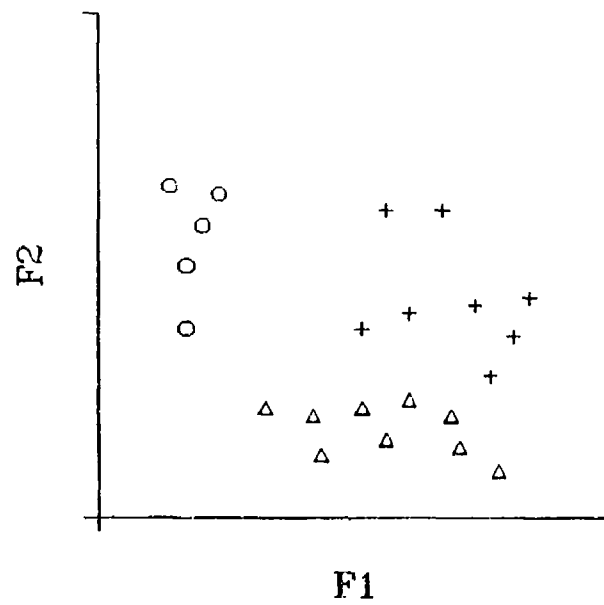
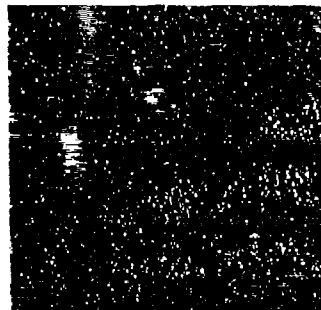


FIGURE III-2. Hypothetical data set in a two-dimensional feature space. Three classes of vectors are represented by different symbols. The top plot shows an effective feature space, the bottom plot an ineffective space.

first-year ice



second-year ice



multiyear ice

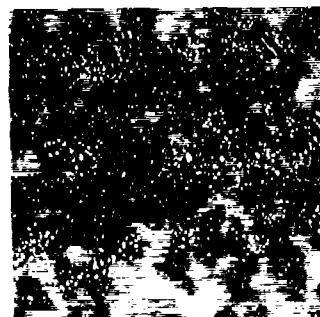


FIGURE III-3. Representative samples of KRMS imagery of first-year, second-year, and multiyear ice.

obviously not be possible to say with certainty from which image the small image fragment was taken. The difference between second-year and multiyear ice types in Figure III-3 is not some unique neighborhood signature but, rather, the average number of texture features per unit area over the entire sample. Therefore, one would expect ice-type classification accuracy to improve as the size of the neighborhood increased because the larger neighborhood gives a better statistical sampling of the texture per unit area. If larger neighborhoods provide more accurate texture features, then why not make neighborhoods extremely large? The limiting factor is that as neighborhoods increase in size, the neighborhoods no longer contain a single ice type and classifications become inaccurate because of ice-type mixtures. Thus, one is faced with the conflicting requirements for large neighborhoods for improved statistics and small neighborhoods to insure single ice types within the neighborhood.

Other investigators have discussed these conflicting requirements for neighborhood size. Connors *et al.* (1984) encountered this problem in the texture-based segmentation of aerial photography of urban areas. They found in their case that an intermediate-sized neighborhood of 145×145 pixels was a good compromise. Weszka *et al.* (1976) found a 64×64 pixel neighborhood size worked well for the classification of terrain types in Landsat Multispectral Scanner imagery. Another solution to this conflict can be found in linear unmixing theory combined with a global approach to image texture analysis. Under the global approach the neighborhood is expanded to include the entire image, thus yielding the best possible estimate of image texture statistics. Linear unmixing theory is then applied to the global estimates of image texture measures to determine the proportions of the various ice types that must be present in the image in order to yield the observed global image texture feature statistics.

In summary, then, the local approach is to determine texture class for each pixel based on its neighborhood, and then to count the pixels in each category to determine the percentage of the image covered by a certain texture type. By contrast, under the global approach, one calculates texture measures based on the entire image, then applies linear unmixing calculations to the global texture statistics to find texture proportions in the image. In both cases the desired result is the same; i.e., proportions of the image covered by each of several texture classes, but the approach to achieve the result is different for the local and global cases. This work follows the global approach.

B. DESIRABLE PROPERTIES OF TEXTURE MEASURES

Figure III-2 showed how effective features should cluster in feature space according to ice type. However, the selection of effective features is more involved than a simple clustering analysis. It is instructive to consider in more depth the question, "what texture measures will be effective features"? What properties are desirable for texture measures? How important are these properties? What are the consequences to the proposed texture analysis if these properties are not realized? How will the presence or absence of these properties be quantitatively measured? Answers to these and related questions form the substance of this section. The properties of texture measures that will contribute to the formation of an effective feature space are enumerated. Each is discussed with the above questions in mind.

1. Correlation with Ice Type

Obviously for a texture measure to be useful in the present application, it must be correlated to some degree with the ice type imaged by the KRMS

scanner. This property has some significant implications. If the texture measures selected for this study are all at least moderately correlated with ice type, then the texture measures will, as a consequence, be correlated to a significant degree with each other. This expected correlation between variables means that the true dimensionality of the multivariate data set will be considerably less than the number of texture measures, or features, defining the feature space. A dimensionality reduction procedure, such as principal components analysis or factor analysis, is therefore indicated as a logical inclusion in the analysis.

How is correlation with ice type to be quantitatively judged? In this work correlation of candidate texture measures with ice type will be determined by performing a classification of fifty 64×64 pixel samples of pure ice types into first-year, second-year, and multiyear classes based on the texture measure in question. (These 50 samples, called data set A, are discussed in Section IV.) The classification algorithm employed will be the maximum Bayesian classifier (Duda and Hart, 1973), which is discussed in Appendix A. The end result of this classification is a confusion matrix from which an overall classification accuracy can be calculated (see Appendix A).

2. Multivariate Normality

Although the linear unmixing algorithms (Sect. VI) do not explicitly require multivariate normality, it is desirable that this property exist in the texture feature space because several ancillary procedures associated with the study do require normality. For example, the Bayesian classifier, just mentioned in connection with the determination of correlation of texture measures with ice type, is based on the assumption of normality. Therefore, the candidate texture measures will be examined for normality. Once normality is established, we can

proceed into various statistical analyses as deemed necessary in the course of the study without being hindered by concerns about normality.

Since texture analysis using the feature space paradigm from statistical pattern recognition is a multivariate problem, the multivariate normality of the data set is the property actually desired. However, the construction of "good" overall tests for joint normality of multivariate data sets has proven to be difficult for the statistical community (Johnson and Wichern, 1982). Therefore, practicality dictates that for the present data set, examination of normality must be limited to the univariate distributions of the individual texture measures. The drawback of examining univariate distributions only is that we can never be certain that we have not missed some nonnormal feature that is only revealed in higher dimensions. The chances of this happening are apparently small, since Johnson and Wichern (1982) state the following.

Fortunately, pathological data sets that are normal in lower dimensional representations but nonnormal in higher dimensions have not frequently been detected.

As a simple means of examining univariate normality, one could create a histogram of the values of the texture measure in question and examine the histogram subjectively for symmetry and for a general "bell-shaped" appearance. To be quantitative, skew and kurtosis could be calculated to give numerical values to histogram symmetry and "peakedness," respectively. However, more sophisticated tests are also available. For example, a common technique for assessing univariate normality is the Q-Q plot, which is a plot of the sample quantile versus the quantile one would expect to observe if the observations actually were normally distributed. Appendix B discusses the calculation of Q-Q plots and gives an example. For a normal distribution the Q-Q plot will be linear. Departures from linear are therefore a measure of the nonnormality of a

data set. The linear correlation coefficient of the Q-Q points will be used here as the criterion of normality.

Several of the texture measures under consideration will be shown to be normal by the linear Q-Q plot test in Section IV. However, a mathematical transformation of variables will be used in these cases to bring the data sufficiently close to constant normality. Precise normality is not a critical factor for this study.

3. Linear Mixing

The approach to texture analysis set forth here requires that image compositions be determined by linear unmixing of global estimates of texture variables. For this approach to be useful, the texture measures selected must mix in a known manner, which for the present study we require to be linear. By mixing linearly we mean that if texture measure x_i has a value $x_i^{(I)}$ for a pure sample from Class I and a value of $x_i^{(II)}$ for a pure sample from Class II, then the value of x_i for a half-and-half mixture of Classes I and II, $x_i^{(I+II)}$, would be

$$x_i^{(I+II)} = \frac{x_i^{(I)} + x_i^{(II)}}{2}. \quad (III-1)$$

To assess the mixing performance of candidate texture measures, we will take typical first-year and multiyear ice-image samples and mosaick them to form a mixture image containing equal proportions of first-year and multiyear ice. Values of the texture measure in question will be calculated for the individual first-year and multiyear samples, as well as for the combined mixture image. Based on the values of a given texture measure calculated from the mosaicked first-year plus multiyear image, the fraction of the combined image occupied by multiyear ice will be linearly predicted. For a texture variable to be considered to mix linearly, the proportion of the image linearly predicted (Eq. III-1))

to contain multiyear ice must fall between 0.4 and 0.6. (The true value is, of course, 0.5.) The ± 0.1 error tolerance is allowed to account for imprecision in texture measure calculations that result from edge effects or other factors that are different between single images and a mosaicked mixture. (Edge effects are discussed further in Section IV.)

The linear mixing property for texture measures is absolutely essential in the present study because the global approach to texture analysis depends on the ability to unmix globally derived texture data. As the study proceeded, the linear mixing requirement proved to be a major roadblock. Section IV will show that only 6 of the 25 texture variables considered behaved linearly in mixtures. Thus, dealing with nonlinear mixing behavior became one of the major study areas in this project.

a. Constant Sum Vectors

The CABFAC and QMODEL codes (see for example, Imbrie and Van Andel, 1964; Kluvan and Pöschel, 1976; Full *et al.*, 1981; Full *et al.*, 1982), which are the best-known computer implementations of the mathematics of linear unmixing, require that the data be "constant sum." The term, constant sum, means that all feature vectors, $\vec{x} = (x_1, x_2, \dots, x_m)$, have components, x_i , such that

$$\sum_{i=1}^m x_i = K \quad (III-2)$$

where K is a constant. Some data sets are naturally constant sum. For example, in geology, rock composition data sum to 100% for every sample. However, the feature vectors here have components that consist of various image texture measures, so they do not necessarily obey Eq. (III-2). If texture measures could be found that naturally formed constant sum vectors, then it would be

an advantage to the unmixing analysis employed here. Such natural constant sum texture variables are not out of the question. For example, if one would calculate the energy spectrum of an image and then use the percent of total energy contained in spectral "bins" as texture measures, the result would be constant sum feature vectors, since the sum of the energy in all bins would be 100% for every sample.

Although some texture measures, such as the energy binning just proposed, may be naturally constant sum, the majority of the texture measures commonly used are not. If the constant sum requirement could be circumvented, then this study would not need to be limited to the small subset of texture measures that provide natural constant sums. One could, of course, scale data vectors so that the sum of the components is forced to the desired sum. However, in doing so the ice-type information in the data set may be lost. Thus, a careful look at the constant sum problem is required. Section V contains an in-depth treatment of this topic.

5. Invariance Under Linear Transformation

A desirable but nonessential characteristic of image texture measures is that they be invariant under linear transformations of the image intensity values. This property is desirable because it would result in texture analysis algorithms that are independent of sensor calibration or scene illumination. Without invariance to linear transformations of image intensities, the form of texture algorithms may be universal from one data set to another, but the numeric values of the coefficients or parameters will change for every data set. In that case, one is faced with the requirement to "train" the texture analysis algorithms on every data set, or possibly even "retrain" during a single experiment to compensate

for drift in sensor calibration. Environmental factors such as atmospheric haze can also introduce linear gain and offset factors into recorded image intensities, requiring "retraining" to compensate for environmental variability. Obviously, the invariance characteristic is highly desirable in general, but texture analysis can still be effectively performed in its absence. A few of the texture measures examined in this study will be shown to be invariant; the majority are not.

The invariance of some of the texture measures could be proven analytically. In other cases the proof must be empirical by means of calculations from actual images before and after a linear transformation is performed. Since some texture measures require examination by empirical means, the analytical proofs have been ignored and all candidate texture measures have been evaluated for invariance based on calculations from a test image. The invariance test was conducted using a multiyear sample from data set A (see Section IV). Let $I(s, l)$ represent the intensity of a test image pixel taken from a location indicated by sample number s and line number l . A new, linearly transformed test image with intensity values $I'(s, l)$, formed from $I(s, l)$ according to

$$I'(s, l) = aI(s, l) + b, \quad (III - 3)$$

where a is the gain and b is the offset values of a linear transformation. For the test of invariance applied to data set A in Section IV, the values $a = 0.8$ and $b = -20$ were used. The texture measure in question was calculated for the original image and also for the transformed image. If the two values agreed within the limits of precision on the calculation, then the texture measure is labeled as invariant.

So few of the texture measures examined were invariant that the noninvariant measures could not be eliminated from consideration in this study. The

information on invariance of candidate texture measures is therefore included in Section IV only as useful information to the reader who may be considering these texture measures in some other context than the present study.

Since all data analyzed here were obtained during a single flight and since the KRMS has proven to be very stable in calibration over the duration of a flight (Eppler, 1987), the invariance properties of texture measures are not of concern here. Such techniques as histogram equalization, which are commonly applied to remove calibration or illumination variability prior to texture analysis (Haralick *et al.*, 1973), were not applied to this data set. The reader is cautioned that numerical results from this experiment can not be transferred directly to other data sets.

6. Computational Simplicity

Computational simplicity is always a desirable property for texture measures. In applications of texture analysis to large data sets or in real-time applications, computational simplicity may emerge as the most desirable of all properties. However, in the present research work with small data sets, computational simplicity is of no consequence. This factor is therefore ignored.

C. CANDIDATE TEXTURE MEASURES

Twenty-five texture measures were considered for inclusion in this study. These measures fell into the following general classes.

- first-order statistics
- measures derived from the gray level co-occurrence matrix (second-order statistics)

- Laws' masks (energy spectrum partitioning)
- gradient analysis
- run length

Representative image texture measures were selected from each of these five groups for a total of 25 texture measures to serve as candidates for inclusion in this ice classification study. This section will describe the candidate texture measures.

The candidate texture measures examined here were selected because they appeared to be the measures most frequently appearing in the literature or because they seemed especially appropriate for sea ice textures. However, numerous texture measures not considered in the present study have also been reported in the literature. Some examples are given below of work on texture measures that have been reported in the literature but not included here.

Texture signatures have been generated using gray-level generalizations of binary "shrink" and "expand" operators (Werman and Peleg, 1985). Harwood *et al.* (1985) studied rank correlation between a standard "rank mask" and a "ranked local neighborhood." This method is similar to convolution with Laws masks, except that the mask and the local neighborhood are converted to ranked form. Ranking is claimed to result in a more robust texture analysis because local order statistics are unaffected by local sample differences due to monotonic shifts of texture gray values and are less sensitive to noise. Peet and Sahota (1985) have drawn upon classical differential geometry by considering the intensity values of an image to form a three-dimensional surface. Measures of local surface curvature, such as the Gaussian curvature or the mean curvature, are applied to texture discrimination. Davis *et al.* (1979) analyze texture by

replacing the texture with another image that indicates the positions of certain local texture properties, e.g., of edges in the original image. Local maxima are then located in the transformed image and a generalized co-occurrence matrix is defined that will capture the important spatial properties of these local maxima. Davis *et al.* (1979) concluded that generalized co-occurrence matrices are useful in distinguishing macro textures that are not satisfactorily distinguishable using features derived from conventional co-occurrence matrices. These examples, and many other texture measures not mentioned here, may be useful for ice classification, but it was necessary in this first work on unmixing global texture measurements to limit the candidate texture measures to a manageable number. The 25 texture measures considered are believed to be representative, but they are not necessarily an optimal selection.

1. First-Order Statistics

The first-order statistical quantities, mean, variance, skew, and kurtosis, will be considered as possible texture measures. These distributional parameters are not normally considered good texture measures because first-order statistics are invariant to rearrangement of pixel locations. An image with given first-order statistics can be scrambled randomly or by design into any possible spatial arrangement of pixels and the values of the first-order statistics remain unchanged. Since texture is "the local spatial distribution of tonal values within an image" (Haralick, 1979), texture measures that don't reflect differences in local spatial distributions are not strictly texture measures. Nevertheless, first-order statistics are often used in texture analysis (e.g., Laws, 1980).

Holmes *et al.* (1984) and Lyden *et al.* (1984) specifically use first-order statistics in connection with ice-type classification in synthetic aperture radar

(SAR) imagery. Lyden *et al.* (1984) observed that for X-band SAR with horizontal polarization on both transmitter and receiver, the image statistics for first-year ice may be distributed differently than for multiyear ice. Image intensity values for first-year ice are approximated by a Rayleigh distribution, but the multiyear ice backscatter is more nearly normal. This suggests that ice types could be classified by means of their distributional characteristics. Distribution differences observed in radar images (backscatter) do not necessarily appear in KRMS images (emissive). However, because of Kirchoff's Law, which relates reflectance to absorptance, enough similarity between active and passive sea ice imagery would be expected to justify including these first-order statistics in the present study. The study will subsequently show that one of these first-order statistics, mean intensity value (or mean brightness temperature), is essential to distinguishing second-year ice from a half-and-half mixture of first-year and multiyear ice.

The mean, μ , of a sample is given by (Dixon and Massey, 1983)

$$\mu = \frac{1}{SL} \sum_{s=1}^S \sum_{l=1}^L I(s, l), \quad (III - 4)$$

where S and L are the total number of samples and lines in the image. If we define μ_i to be the i^{th} central moment about the mean, μ_i is given by (Bury, 1975)

$$\mu_i = \frac{1}{SL} \sum_{s=1}^S \sum_{l=1}^L (I(s, l) - \mu)^i. \quad (III - 5)$$

Of course, the first central moment, μ_1 is always zero. For $i = 2$, μ_2 reduces to the well-known equation for sample variance (Dixon and Massey, 1983). Skew, a measure of the symmetry of the distribution, and kurtosis, a measure of peakedness of the distribution, are calculated from the third and fourth central moments

as follows.

$$skew = \frac{\mu_3}{\mu_2^{3/2}} \quad (III - 6)$$

$$kurtosis = \frac{\mu_4}{\mu_2^2} \quad (III - 7)$$

The mean, the variance, and the kurtosis always assume positive values for image data, where $I(s, l)$ is always positive. Skew can assume either positive or negative values. To avoid dealing with negative numbers, skew values here have been offset by +3. Therefore, what we call skew is actually $(skew + 3)$. Texture variables derived from these first-order statistics will be called MEAN, VAR, SKEW, and KURT from this point forward.

2. Gray Level Co-occurrence Matrix

The gray-level co-occurrence (GLC) matrix, $P_{i,j}$, is an array of second-order probabilities. The i, j element is the probability of finding a pixel with gray level j at some given spatial offset, $\vec{\delta}$, from a pixel with gray-level intensity value i . A number of GLC matrices can be calculated for a given image. Each possible value of $\vec{\delta}$ gives rise to an individual GLC matrix. In digital imagery we can express the displacement as a two-component vector, where one component in the samples (columns) direction has a magnitude of $m\Delta s$, and one component in the lines (rows) direction has a magnitude of $n\Delta l$.

$$\vec{\delta} = (m\Delta s, n\Delta l), \quad (III - 8)$$

where m and n are integers and Δs and Δl are the sampling intervals in the samples and lines directions, respectively. It is therefore convenient to consider the GLC matrix to be a four-dimensional array, $P_{i,j,m,n}$.

GLC matrices contain various types of texture-related information. For example, the GLC matrix obviously contains edge information. If an image is

uniform, the i and j intensity values at the ends of the displacement vector will always be nearly equal, thus contributing to larger probabilities in the near-diagonal elements of the GLC matrix. Conversely, displacement vectors spanning edges link i, j values that are significantly different, thus increasing the values of the probabilities in the off-diagonal elements of the GLC matrix. The ratio of the values of the near-diagonal to off-diagonal elements of the matrix therefore are a measure of "edginess."

The GLC matrix also contains first-order statistics. If the rows or columns of the GLC matrix are summed, then the result is the histogram of the image from which all of the first-order statistics discussed could be derived. Texture measures derived from the GLC matrix are, therefore, redundant to some degree with the first-order statistics.

Shape information is also contained in the GLC matrix. To illustrate this fact, consider the binary images of the letters T and L shown in Figure III-4, along with their corresponding GLC matrices for $\vec{\delta} = (-\Delta s, \Delta l)$. Note that these two images have identical histograms (first-order statistics) and the same number of edge points, yet the GLC matrices are different. This difference between GLC matrices for the letters T and L is attributed by Trivedi and Harlow (1985) to the fact that the GLC matrices contain at least some rudimentary shape information.

With first-order statistics, edge information, and shape information influencing the distribution of second-order probabilities within the GLC matrix, the GLC matrix is obviously affected by the local distribution of gray shades within an image, and therefore offers promise for texture analysis. Second-order probabilities have a long history in association with texture analysis. Physiologists studying the human visual system suggested long ago that second-order

gray level probabilities are what the human vision system is actually detecting in visual perception of texture. Julesz (1962) stated what subsequently became known as the Julesz Conjecture:

Second-order gray level probabilities could match at least the primitive level of human texture perception.

Stating the Julesz Conjecture in more mathematical terms, we would say that a necessary condition for two textures to be visually discriminable is that they have different second-order probabilities.

The Julesz Conjecture stood for several years, but eventually papers appeared presenting counterexamples to this conjecture. That is, textures were synthesized that had identical second-order probabilities, yet were visually discernible. Some of the early counterexamples were not very convincing, since they were barely discernible to the human visual system. However, quite distinct examples were eventually found. Figure III-5 is an example of two visually distinct textures that have identical second-order probabilities.

Gagalowicz (1979, 1981) salvaged the Julesz Conjecture by announcing a revision. Harlow and Connors (1983) state this revision as follows:

A necessary condition for two textures to be discriminable is that they have different local second order probabilities computed over a small region of the image.

This revision, which specified local second-order probabilities, has served to maintain the idea that the GLC matrix contains sufficient information to allow close match with the human level of texture perception. Therefore, the texture measures derived from the GLC matrix appear to be the most widely used texture measures. The GLC matrices themselves are very unwieldy. For 8-bit digital imagery, where the number of possible gray levels, L , is 256, an individual GLC matrix would be dimensioned 256×256 . It would be difficult to find texture-

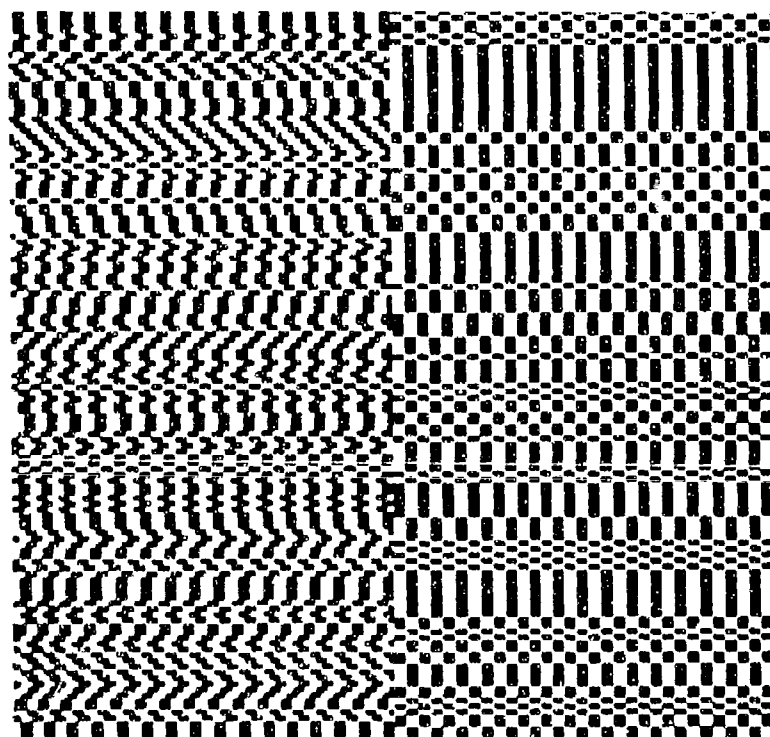


FIGURE III-5. Two visually distinct textures that have identical second-order probabilities. (Taken from Harlow and Conners, 1983.)

related differences in the values of these 65,536 individual matrix elements. The customary approach is, therefore, to work with "texture measures" derived from the GLC matrix rather than with the matrix elements themselves. Haralick *et al.* (1973) wrote the classical paper defining texture measures derived from the GLC matrix. He called the GLC matrix the Gray Tone Spatial Dependence Matrix. This paper defined 14 texture measures calculated by various weighted combinations of the GLC matrix elements. Some of these texture measures have found widespread use; others appear to have never been used after the original paper. Other authors, such as Connors (1979), have proposed additional texture measures from the GLC matrix. The cluster prominence measure that follows is one measure that came from Connors' (1979) work.

Six of the most widely accepted GLC matrix-based texture measures have been included in this study and are listed below, along with their definitions.

$$\begin{aligned}
 inertia(m, n) &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - j)^2 P_{i,j,m,n} \\
 cluster\ prominence(m, n) &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i + j - \mu_i - \mu_j)^4 P_{i,j,m,n} \\
 local\ homogeneity(m, n) &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{1}{(1 + (i - j)^2)} P_{i,j,m,n} \\
 energy(m, n) &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_{i,j,m,n}^2 \\
 entropy(m, n) &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} P_{i,j,m,n} \ln(P_{i,j,m,n}) \\
 correlation(m, n) &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} \frac{(i - \mu_i)(j - \mu_j) P_{i,j,m,n}}{\sigma_i \sigma_j},
 \end{aligned}
 \tag{III -- 9}$$

where

$$\begin{aligned}\mu_i &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} i P_{i,j,m,n} \\ \mu_j &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} j P_{i,j,m,n} \\ \sigma_i^2 &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (i - \mu_i)^2 P_{i,j,m,n} \\ \sigma_j^2 &= \sum_{i=0}^{L-1} \sum_{j=0}^{L-1} (j - \mu_j)^2 P_{i,j,m,n}.\end{aligned}$$

Naming of these texture features is not consistent between authors. For example, what we call inertia and local homogeneity, Haralick *et al.* (1973) called contrast and inverse difference moment, respectively. The names used here are consistent with more recent work, such as Ballard and Brown (1982) and Trivedi and Harlow (1985).

In addition to the selection of these six texture measures, the displacement vectors to be used for calculation of GLC matrices also must be considered. For some texture measures, maximum ice-type discrimination occurs using large displacement vectors; for other texture measures, small displacements give better results. Figure III-6 shows the inertia measure for three representative samples of first-year (FY), second-year (SY), and multiyear (MY) ice as a function of the length of $\vec{\delta}(m,n)$. In this case maximum ice-type separation occurs for large displacement lengths. Figure III-7 shows the correlation measure as a function of displacement vector length for the same three representative samples used for Figure III-6. In the case of the correlation measure the best separation of ice-types results at small displacement vector lengths. To cover both of these cases, we have utilized both short and long displacement vector lengths in this study.

Note in Figures III-6 and III-7 that several of the curves have a ripple

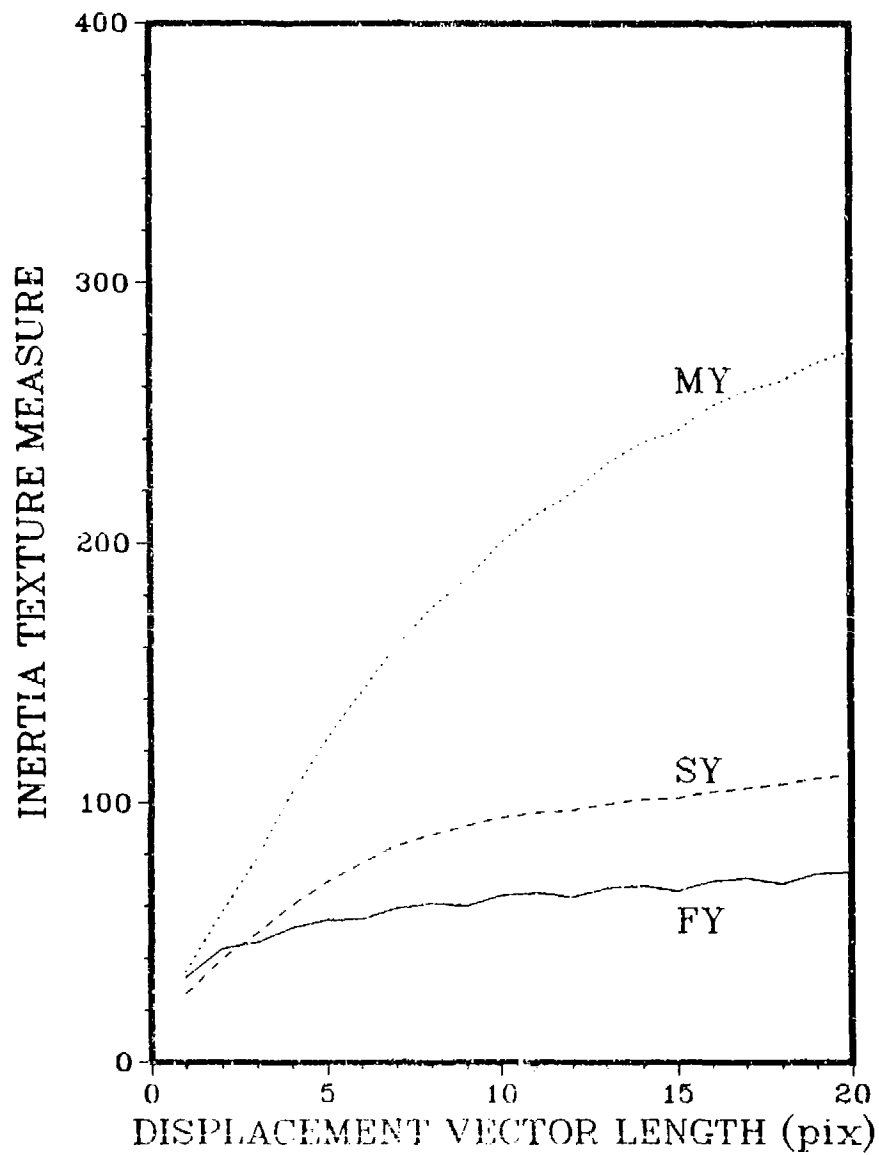


FIGURE III-6. The GLC matrix-based inertia texture measure as a function of displacement vector length for three typical 128 x 128 pixel samples of FY, SY, and MY ice.

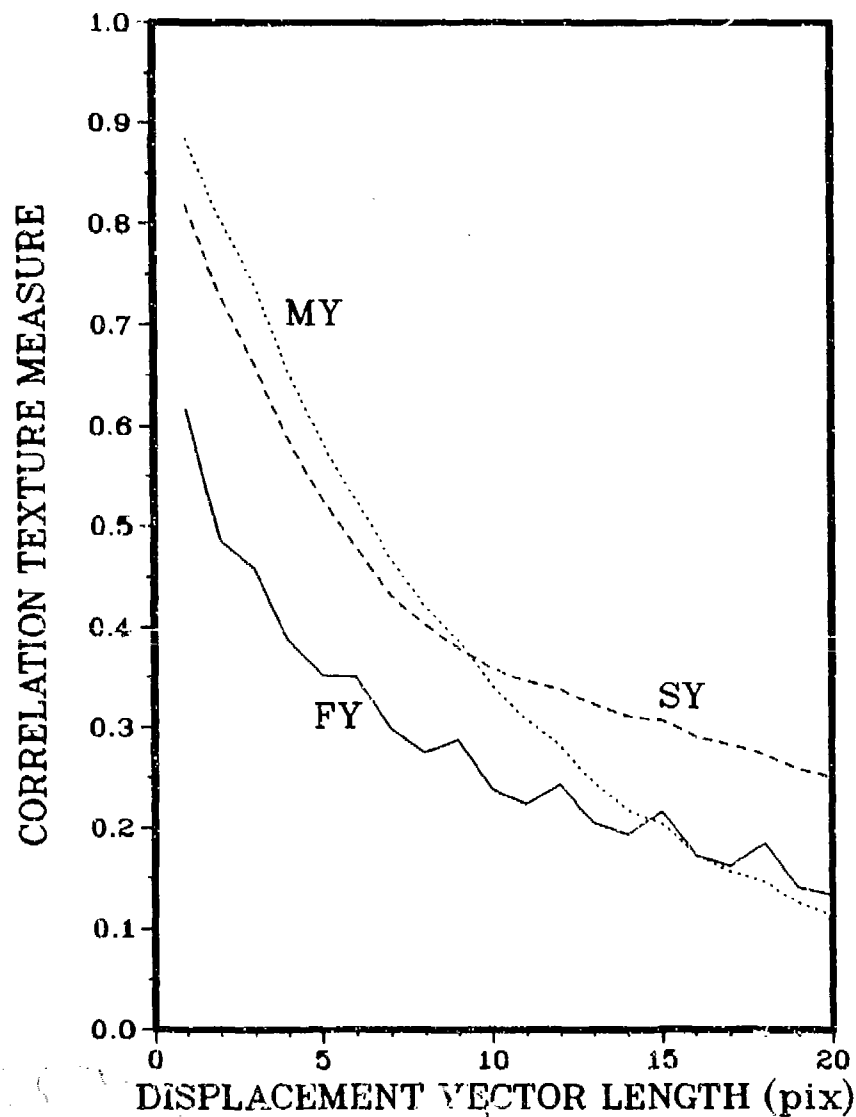


FIGURE III-7. The GLC matrix-based correlation texture measure as a function of displacement vector length for three typical 128 x 128 pixel samples of FY, SY, and MY ice.

with a period of three pixels. The ripple is the result of imbalance between the three antennas in the KRMS scanner. For displacement vectors with along-track components, which are multiples of three, both ends of the displacement vector cover pixels sensed by the same antenna. For along-track displacement components, which are not multiples of three, the two pixels located at the ends of the vector are sensed by different antennas, in which case the texture information is corrupted by a contribution from antenna-to-antenna imbalance. It is obviously better, then, to use only displacement vector lengths that are multiples of three in the along-track direction. For this reason we have selected displacement vector lengths of 3 and 15 pixels to serve as the short and long displacement values, respectively. Each of the six GLC matrix-based texture measures was calculated for both of these displacements. The result is a total of 12 texture measures based on GLC matrix calculations.

For a given displacement vector length two GLC matrices were calculated, one with $\vec{\delta}$ in the samples direction and the other with $\vec{\delta}$ in the lines direction. The value of the texture measure utilized in subsequent texture analysis is the average of these two values. For example the inertia texture measure for displacement vector length, k is given by

$$inertia(k) = \frac{inertia(k\Delta s, 0) + inertia(0, k\Delta l)}{2}. \quad (III - 10)$$

Although texture measures derived from the second-order probabilities contained within the GLC matrix are widely used for texture analysis, the only previous work known to the author where these texture measures were applied to ice-type discrimination is that of Holmes *et al.* (1984). This previous study involved the use of the entropy and inertia measures to classify sea-ice types in SAR imagery.

The GLC matrix-based texture measures listed in Eq. (III-9) will be given the variable names INR3, PR3, HOM3, ENG3, ENT3, and COR3 when calculated with a displacement vector length of three pixels. The names INR15, PR15, HOM15, ENG15, ENT15, and COR15 will be used when calculations utilize a 15-pixel displacement vector length.

3. Laws' Masks

Laws (1980) proposed a set of spatial domain convolution filters, the output of which he used to classify an image texture data set with better accuracy than could be achieved with GLC-based texture measures. The Laws method will be shown as simply an image domain implementation of the old idea of power spectrum partitioning. Laws' approach to energy partitioning will be followed here for the calculation of image texture measures related to the frequency distribution of image energy.

If texture is the local distribution of intensity values within an image, then differences of texture will clearly be reflected in the two-dimensional power spectrum of the image. To look at spectral distribution differences, some investigators work with the power spectrum while others work with the autocorrelation function of the image. These approaches are equivalent, since the autocorrelation function and the power spectrum form a Fourier transform pair according to the well-known autocorrelation theorem (Bracewell, 1965).

Of those who work in the spatial-frequency domain, some search the Fourier space directly for features in the power spectrum that can be related to image texture. For example, Bajscy and Lieberman (1974, 1976) determined that blob-like textures tend to have peaks in the power spectrum at radii comparable to the size of the blobs. A more common approach to spatial-frequency

content analysis is to partition Fourier space into bins. Two kinds of bins, radial and angular, are often used. It is well known that the radial distribution of values in the power spectrum is sensitive to texture coarseness. A coarse texture will result in high values concentrated near the origin of the spectrum, but a fine texture generates an energy spectrum that will be more spread out. Similarly, it is well known that the angular distribution of values in the power spectrum is sensitive to directionality of the texture. A texture with many edges or lines in a given direction, θ , will have high spectral energy values concentrated around the direction $\theta + (\pi/2)$ from the origin of the spectrum. Thus, a good set of texture measures reflecting coarseness and directionality could be expected from measuring the energy content in a set of radial and/or angular bins within the Fourier power spectrum of the image. Lendaris and Stanley (1969; 1970) used Fourier spectrum partitioning to discriminate between natural and man-made scenes in low-altitude photographs. Weszka *et al.* (1976) also apply Fourier power spectral analysis to terrain classification.

In general, texture features based on Fourier power spectra are considered to perform more poorly than features based on second-order, gray-level co-occurrence statistics (Haralick, 1979). However, Laws (1980) reports that his "texture energy" method performed better than GLC matrix-based texture features on the Brodatz (1966) texture set he used. This study will show that the texture energy measures generated by Laws' masks yield nearly equivalent performance to the GLC texture measures here in the case of ice textures.

Laws (1980) proposed the three sets of one-dimensional convolution masks shown in Figure III-8. The logic leading to the selection of these masks is not obvious. These masks do have some interesting characteristics. Each mask is weighted more heavily toward the center; all are either symmetric or antisymmetric,

THIRD-ORDER MASKS:

$$\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 1 \\ -1 & 2 & -1 \end{bmatrix}$$

FIFTH-ORDER MASKS:

$$\begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ -1 & -2 & 0 & 2 & 1 \\ -1 & 0 & 2 & 0 & -1 \\ -1 & 2 & 0 & -2 & 1 \\ 1 & -4 & 6 & -4 & 1 \end{bmatrix}$$

SEVENTH-ORDER MASKS:

$$\begin{bmatrix} 1 & 6 & 15 & 20 & 15 & 6 & 1 \\ -1 & -4 & -5 & 0 & 5 & 4 & 1 \\ -1 & -2 & 1 & 4 & 1 & -2 & -1 \\ -1 & 0 & 3 & 0 & -3 & 0 & 1 \\ 1 & -2 & -1 & 4 & -1 & -2 & 1 \\ 1 & -4 & 5 & 0 & -5 & 4 & -1 \\ -1 & 6 & -15 & 20 & -15 & 6 & -1 \end{bmatrix}$$

Figure III-8. Center-weighted vector masks (Laws, 1980).

and all but the first mask of each order are zero sum. The center-weighted aspect of these masks is ideal for a spatial convolution filter because the tapered ends avoid severe frequency domain sidelobes that result from box-car-truncated, spatial domain convolution filters. Laws states that the vectors formed by the mask coefficients are independent within each order, but are not orthogonal.

The third-order masks form the basis for the higher-order masks. Each fifth-order mask can be generated by convolving two of the third-order masks. The seventh-order masks can be generated by convolving a third-order mask with a fifth-order mask, or by twice convolving third-order masks. Note that the third-order masks contain the polynomial coefficients of the products $(a + b)(a + b)$, $(a + b)(a - b)$, and $(a - b)(a - b)$. Actually, all of the masks can be generated directly from a binomial expansion of an appropriate number of terms rather than by convolution of lower-order masks.

Laws (1980) generated two-dimensional masks by forming vector cross products of the one-dimensional masks. In the present study we have convolved the images twice with a one-dimensional mask rather than form cross-product matrix masks. The vector mask is applied first as a horizontal kernel, and then the output of that operation is convolved with the vector mask applied as a vertical kernel. Double application of vector masks in this fashion produces results identical to a single convolution with a cross-product matrix kernel. The reason for double application of a one-dimensional mask is to avoid the chore of entering all of the elements of a large matrix kernel into the computer. The resulting energy binning is therefore in the form of two-dimensional rings of varying radii in the image energy spectrum. Radii of the rings will be the same as the distance of the peak in the mask Fourier transforms from the origin of the plots in Figures III-9 or III-11.

We now proceed to establish the fact that the variance of the images resulting from convolutions with Laws' masks are actually the values of image energy falling in a series of spectral bins. Consider the seventh-order masks given in Figure III-8. Each of these masks has been Fourier transformed to obtain its frequency domain equivalent. The Fourier transform squared (energy spectrum) for each of the masks is shown in Figure III-9. Note that the spectra associated with these masks are narrow-banded with center frequencies ranging from DC to the Nyquist frequency, f_N , as the mask numbers go from 1 to 7. Referring to the convolution theorem (Bracewell, 1965) we note that the convolution of the image with the mask is equivalent to multiplying the Fourier transform of the image by the Fourier transform of the mask. It is obvious from the narrow-banded nature of the Fourier transforms of the mask that any such frequency-domain product would contain significant non-zero values in only a limited range of frequencies. Thus, convolution with a Laws' mask is equivalent to applying a band-pass filter to the image. The variance of the image output from the convolution operation is, then, the energy in the original image that falls within the bandpass of the convolution filter. Therefore, the Laws' mask texture energy measures are nothing more than a simple spectral binning of the image contents.

In Section II it was established that the oversampling of the KRMS imagery precluded any high-frequency content in the images. We do not expect any useful ice discrimination information at frequencies above $f_N/3$. Therefore, of the 7 energy bins resulting from the seventh-order Laws' masks, only masks 1 and 2 are expected to contain anything other than random noise. A preferable arrangement for the KRMS image data would be to locate our 7 spectral bins to span the range from DC to $f_N/3$. To perform this arrangement, Laws' masks of 21st order can be constructed and only the first seven used. Such masks have

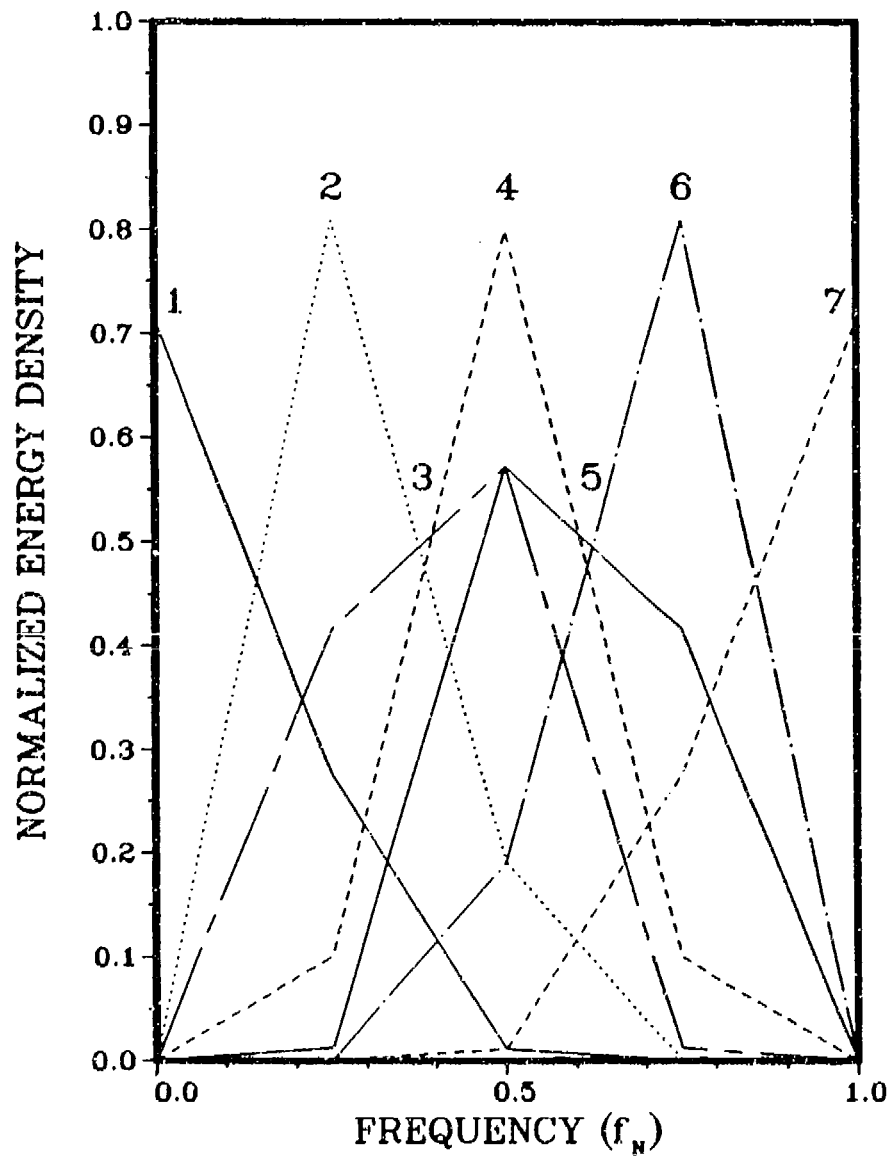


FIGURE III-9. Normalized energy density spectra (Fourier transform squared) of the seventh-order masks shown in Figure III-8.

been constructed by repeated convolutions of the appropriate lower-order masks. Figure III-10 lists the first 7 of the 21st order masks. Energy spectra of these selected 21st order masks are plotted in Figure III-11.

Figure III-12 shows a typical sample of KRMS imagery over multiyear ice and the result of convolving the masks listed in Figure III-10 with this sample image. Each of the convolution results has been individually enhanced to show whatever features may be present in that particular image. Absolute comparisons between the various filter outputs are, therefore, not possible. However, what is obvious is the transition from lower-to-higher frequency content for masks 1 through 7. This transition visually demonstrates the frequency-binning properties that have been attributed to these masks. The "basketweave" texture in the higher-frequency filter outputs appeared as a surprise. A definitive explanation of this appearance is not available. It seems obvious that these patterns are not related to the ice scene imaged by the KRMS, but are artifacts of some type of coherent noise within the KRMS system. It is unfortunate that these higher-frequency bins are dominated by this noise, which would be expected to completely mask any ice texture information that might be available in these spectral bins. However, we will carry all 7 Laws' mask texture energy variables forward as candidate texture measures on the outside chance that some useful ice-type information may actually be present.

The 7 texture energy values associated with the 21st order masks given in Figure III-10 will be carried into the ice-texture analysis as candidate texture measures. These measures will be named LM1 through LM7.

4. Gradient Analysis

Rosenfeld and Thurston (1971) have considered texture in terms of the

21st ORDER MASKS:

1	1	-1	1	1	-1	-1
20	18	-16	14	12	-10	-8
190	152	-118	88	62	-40	-22
1140	798	-528	322	172	-70	-8
4845	2907	-1581	731	237	5	83
15504	7752	-3264	952	-16	248	160
38760	15504	-4488	272	-664	400	-8
77520	23256	-3264	-1496	-1104	40	-352
125970	25194	1326	-3094	-494	-650	-338
167960	16796	7072	-2652	936	-780	208
184756	0	9724	0	1716	0	572
167960	-16796	7072	2652	936	780	208
125970	-25194	1326	3094	-494	650	-338
77520	-23256	-3264	1496	-1104	-40	-352
38760	-15504	-4488	-272	-664	-400	-8
15504	-7752	-3264	-952	-16	-248	160
4845	-2907	-1581	-731	237	-5	83
1140	-798	-528	-322	172	70	-8
190	-152	-118	-88	62	40	-22
20	-18	-16	-14	12	10	-8
1	-1	-1	-1	1	1	-1

Figure III-10. Center-weighted vector masks.

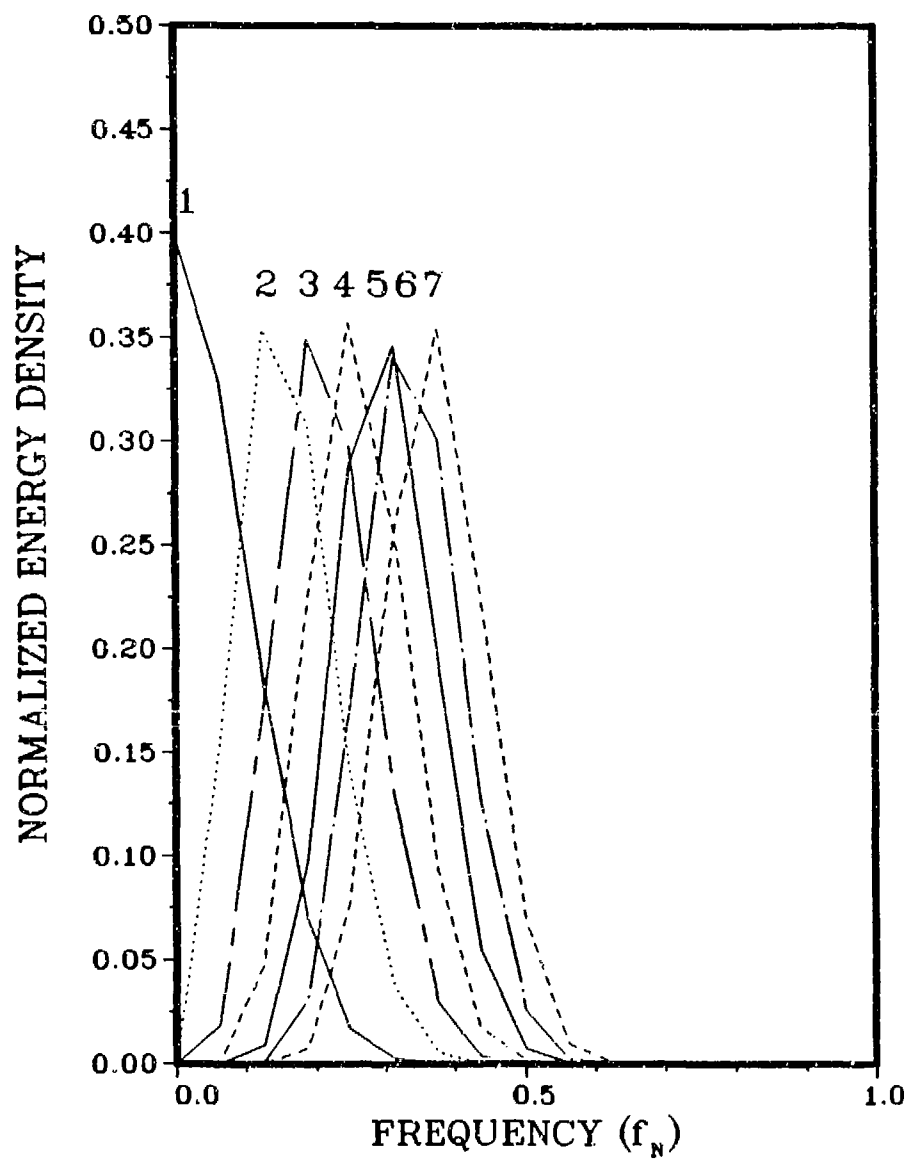


FIGURE III-11. Normalized energy density spectra (Fourier transform squared) of the 21st order masks shown in Figure III-10.

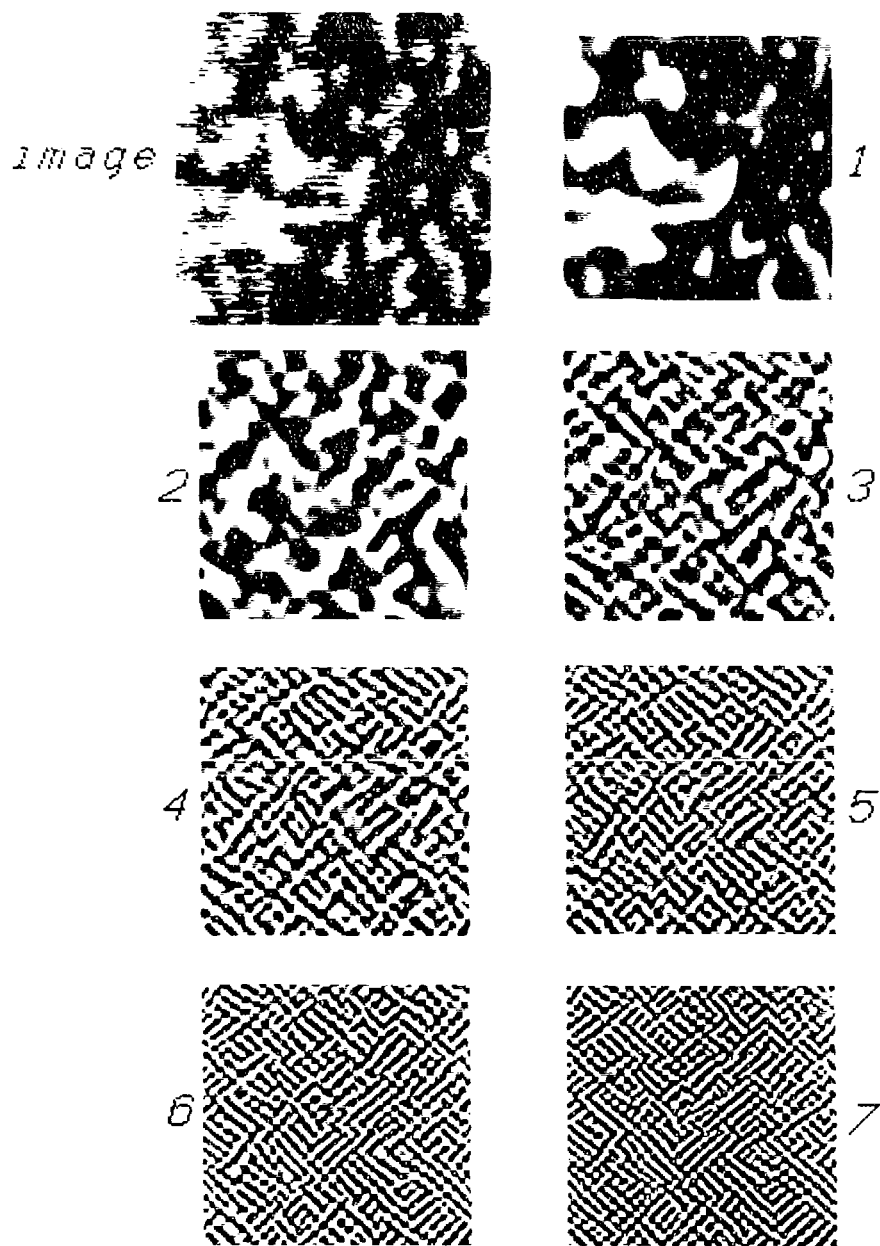


FIGURE III-12. Illustration of image energy spectrum partitioning by spatial convolution of a 128 x 128 pixel KRMS sample of multiyear ice (upper left) with the 21st order masks given in Figure III-10. Images labeled 1 through 7 are the result of the successive horizontal and vertical application of each of the masks to the original image.

amount of "edge" per unit image area. An edge can be detected by a variety of local mathematical operators, which essentially measure some property related to the gradient of the image intensity function. Rosenfeld and Thurston (1971) used the average Roberts gradient over a neighborhood as a measure of texture. Other investigators have used the second derivative of the image intensity function for texture analysis. The second derivative approach is actually the gradient of the gradient. Triendl (1972) used the Laplacian operator to approximate the second derivative of the image and then extracted texture parameters from a smoothed version of the second derivative image.

In addition to the straightforward use of gradients or gradients of gradients, more elaborate forms of gradient analysis have been proposed. Landeweerd and Gelsema (1978) generated textural parameters from the histogram of the gradient image. Rosenfeld (1975) used a double gradient approach. He generated an image whose intensity is proportional to the edge per unit area of the original image. Then the edge density image is further processed by additional gradient analysis prior to textural feature extraction.

Here, a simple edge density approach was chosen to generate a texture measure representing the gradient analysis category of texture measures. An edge detection operator is applied to the image; the result is a gradient image. The mean intensity of the entire gradient image is then taken to be a global measure of the average edge per pixel. The edge operator chosen consists of spatial convolution with an edge detection mask. A number of edge detection convolution kernels have been developed. The mask chosen here is the Prewitt operator (Prewitt, 1970). This operator was chosen because the edges in the KRMS imagery tend to be smooth, ramp-like transitions rather than sharp discontinuities. Nashburg and Lineberry (1981) have shown that the Prewitt

operator maximizes the output signal-to-noise ratio for uniform slope edges corrupted by additive Gaussian noise. It is also less sensitive than the more common Sobel and Roberts operators to noise of any variance (Hayden *et al.*, 1987).

The Prewitt operator consists of a pair of convolution kernels for obtaining the x and y components of the gradient of the image intensity function. These components, $g_x(s, l)$ and $g_y(s, l)$, are obtained via convolution of the image with the following kernels.

$$\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

From the $g_x(s, l)$ and $g_y(s, l)$ images we can form a gradient magnitude image, $g(s, l)$, according to

$$g(s, l) = \sqrt{g_x(s, l)^2 + g_y(s, l)^2}. \quad (III - 11)$$

The mean intensity of the Prewitt gradient magnitude image is incorporated into the global texture analysis under the name EDEN for edge density. Figure III-13 shows the Prewitt edge magnitude images for typical KRMS images of first-year and multiyear ice. The difference in mean brightness of these edge images is apparent.

5. Run Length

The run, defined as a maximal-connected colinear set of pixels, all with the same intensity value, can be a useful texture measure. Various parameters associated with runs could include the length of the run, gray-level value of the run, or angular orientation of the run. Galloway (1975), Weszka *et al.* (1976), Maleson *et al.* (1977), and Werman and Peleg (1985) have used various combinations of these parameters associated with gray-level runs for texture analysis.

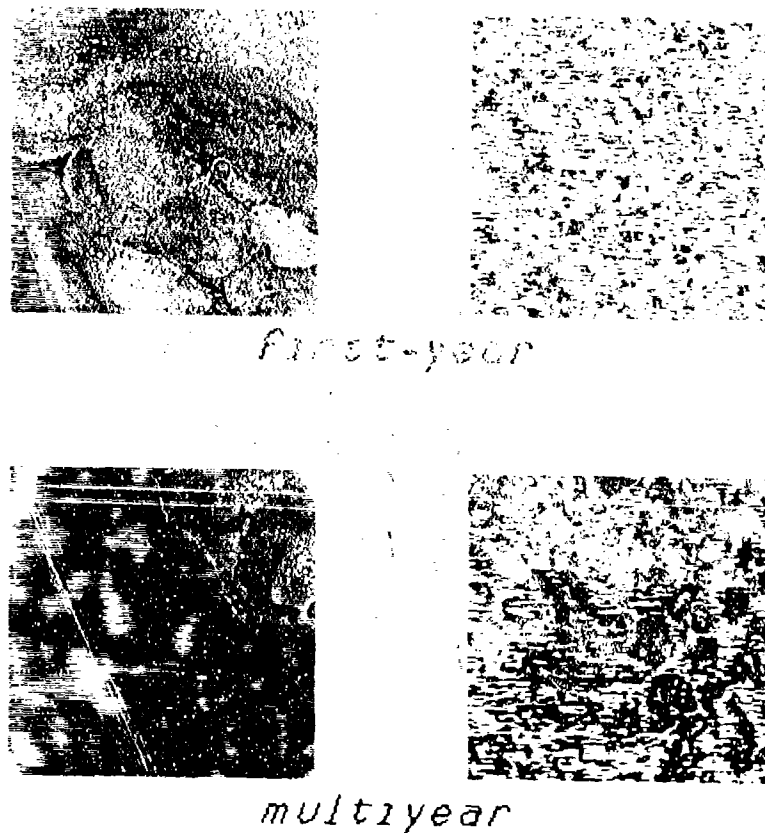
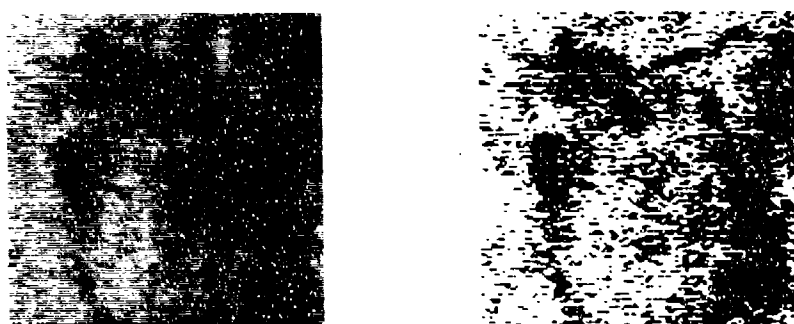


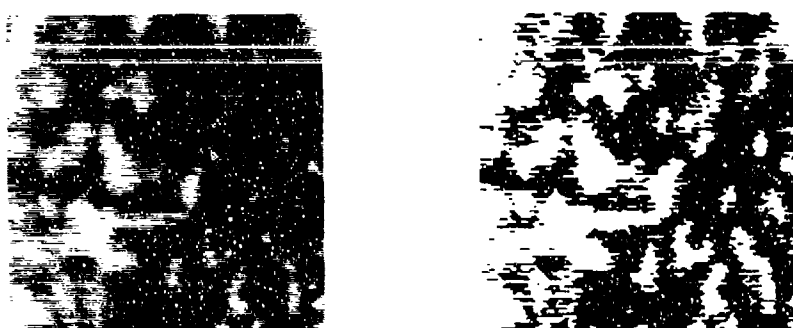
FIGURE III-13. Representative KRMS images of first-year and multi-year ice (left) with their corresponding edge magnitude images (right) formed by convolution of the image with the Prewitt edge operators. Edge magnitude shown here is formed from both the x and y edge components according to Eq. III-11.

Gray-level run parameters are not expected to be good measures of texture for KRMS data because of the high noise levels with respect to scene contrast. In other words, because of the noise in the imagery, adjacent pixels in the image are very seldom at the same intensity level, even if the scene is uniform. Therefore, most gray-level run lengths in the image would be 1, and this length is a measure of the system noise, not the texture in the scene. However, the gray-level image might be converted into a binary image by thresholding the image at its mean value. That is, all pixels with intensity values above the mean are assigned a brightness value of 1 and all pixels with intensity below the mean are assigned a brightness value of 0. In this binary image one could argue that the run lengths might reveal useful ice-type information, as follows.

For first-year ice, which has a very uniform scene brightness, intensity levels would be expected to shift back and forth around the mean on a pixel-to-pixel basis. The expected value of run length in this case would be small, perhaps two pixels or less on the average. However, with multiyear ice, both ridges and melt ponds produce patches of higher or lower intensity in the KRMS imagery. These patches may create areas within the image where all pixels remain above or below the mean brightness temperature, creating uniform patches of 0s and 1s over large areas. The mean run length in this case would be expected to be significantly larger than was the case for the more uniform first-year ice. This expectation is somewhat contrary to intuition, which tells us that the more uniform ice would contain longer run lengths than the patchy ice. However, it is the noise level and the thresholding into a binary image, combined with scene uniformity, that determines mean run length within this type of thresholded image. When these three factors operate in concert, uniform scenes result in shorter binary run lengths than do patchy scenes as shown in Figure III-14.



first-year



multiyear

FIGURE III-14. Representative KRMS images of first-year and multi-year ice (left) with their corresponding binary images (right) formed by thresholding the image at its mean intensity value.

Therefore, the mean binary run length (MBRL) of KRMS images thresholded at their mean value has been included as a candidate texture measure.

Two values of MBRL are calculated for a given image sample. One value is for runs in the samples direction and the other for runs in the lines direction. The MBRL value passed into the texture studies is the arithmetical mean of the values computed individually for the two directions. MBRL values vary considerably in these images as a function of direction. This variability is not believed to be an indication of anisotropy in the image texture but, rather, an artifact resulting from the KRMS antenna-to-antenna imbalance described in Section II. MBRL values, therefore, are consistently lower in the lines direction than in the samples direction. However, in the lines direction MBRL is still correlated with ice type, so that averaging the two directional calculations of MBRL produces useful texture measures for ice-type discrimination.

6. Summary of Candidate Texture Measures

Twenty-five texture measures have been described here as candidates for use in discrimination of sea-ice types. These candidates will be carried forward into the texture analysis, where some will be eliminated after evaluation on ice image data according to the criteria established for texture measures at the beginning of this section. Table III-1 summarizes these candidate texture measures.

TABLE III-1. CANDIDATE IMAGE TEXTURE MEASURES

first-order statistics

mean	MEAN
variance	VAR
skew	SKEW
kurtosis	KURT

second-order statistics (GLC matrix)

inertia	INR3, INR15
cluster prominence	PR3, PR15
local homogeneity	HOM3, HOM15
energy	ENG3, ENG15
entropy	ENT3, ENT15
correlation	COR3, COR15

Laws' masks (spectral binning)

lowest frequency bin	LM1
•	LM2
•	LM3
•	LM4
•	LM5
•	LM6
highest frequency bin	LM7

edge density

mean Prewitt edge magnitude	EDEN
-----------------------------	------

run length

mean binary run length	MBRL
------------------------	------

IV. THE TEXTURE ANALYSIS DATA SET

To determine if the candidate texture measures described in Section III-C possess the properties listed as desirable in Section III-B, a set of fifty 64×64 pixel sample images were extracted from the available KRMS imagery. These 50 samples, referred to henceforth as data set A, are described in this section. Values of the candidate texture measures will be calculated for each of the 50 sample images. Texture measures are then evaluated and discussed according to the criteria established in Section III

A. KRMS SAMPLE IMAGES EXTRACTED AS DATA SET A

Each of the sample images in data set A contain a single ice-type. The ice types were determined by visual interpretation of the KRMS images and coincident aerial photographs (Eppler and Farmer, 1983). First-year, second-year, and multiyear ice are the only ice types contained in the KRMS imagery in adequate amounts to yield reliable texture statistics. The World Meteorological Organization (WMO, 1979) combines these three categories into only two: first-year ice and old ice. First-year ice is ice of not more than one winter's growth that ranges in thickness from 30 cm to 2 m. Old ice is ice that has survived at least one summer's melt and commonly ranges in thickness from 2 to 4 m. For this study the old-ice category has been divided into second year and multiyear. Second-year ice has survived one summer's melt. Multiyear ice has survived more than one summer's melt. The ice imaged by this KRMS mission in March 1983 in the Beaufort Sea offshore from Barrow, Alaska, was predominantly first-year and second-year ice. Multiyear ice was observed very infrequently, so its representation in data set A is limited.

Figures IV-1 through IV-10 show the ten 512×512 pixel KRMS images



FIGURE IV-1. Full 512 x 512 pixel image showing locations where 64 x 64 pixel subareas were extracted as samples MY1 through MY7 and FY5 of data set A.



FIGURE IV--2. Full 512 x 512 pixel image showing locations where 64 x 64 pixel subareas were extracted as samples MY8 and MY9 of data set A.

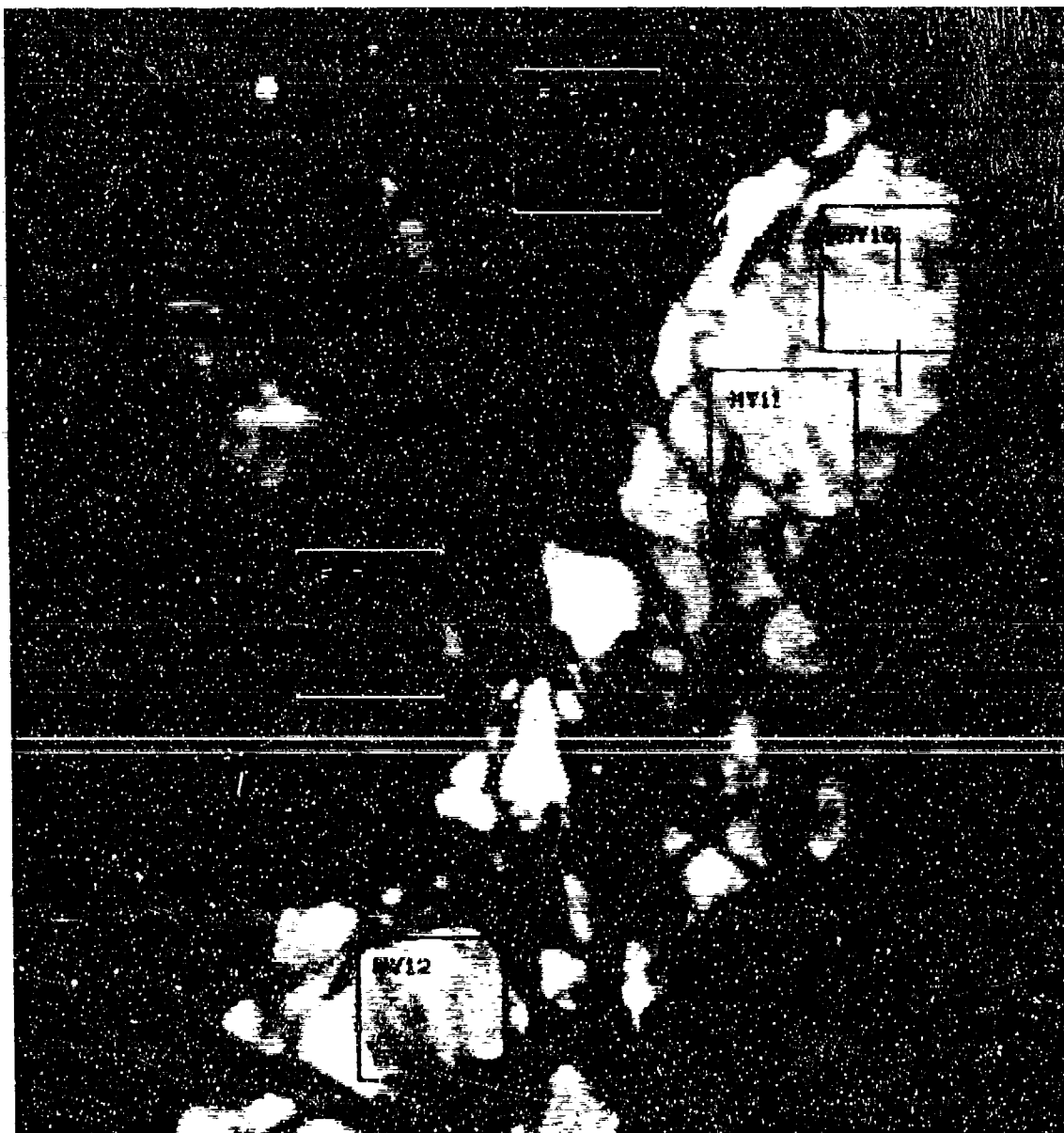


FIGURE IV-3. Full 512 x 512 pixel image showing locations where 64 x 64 pixel subareas were extracted as samples MY10 through MY12 and FY6 through FY7 of data set A.

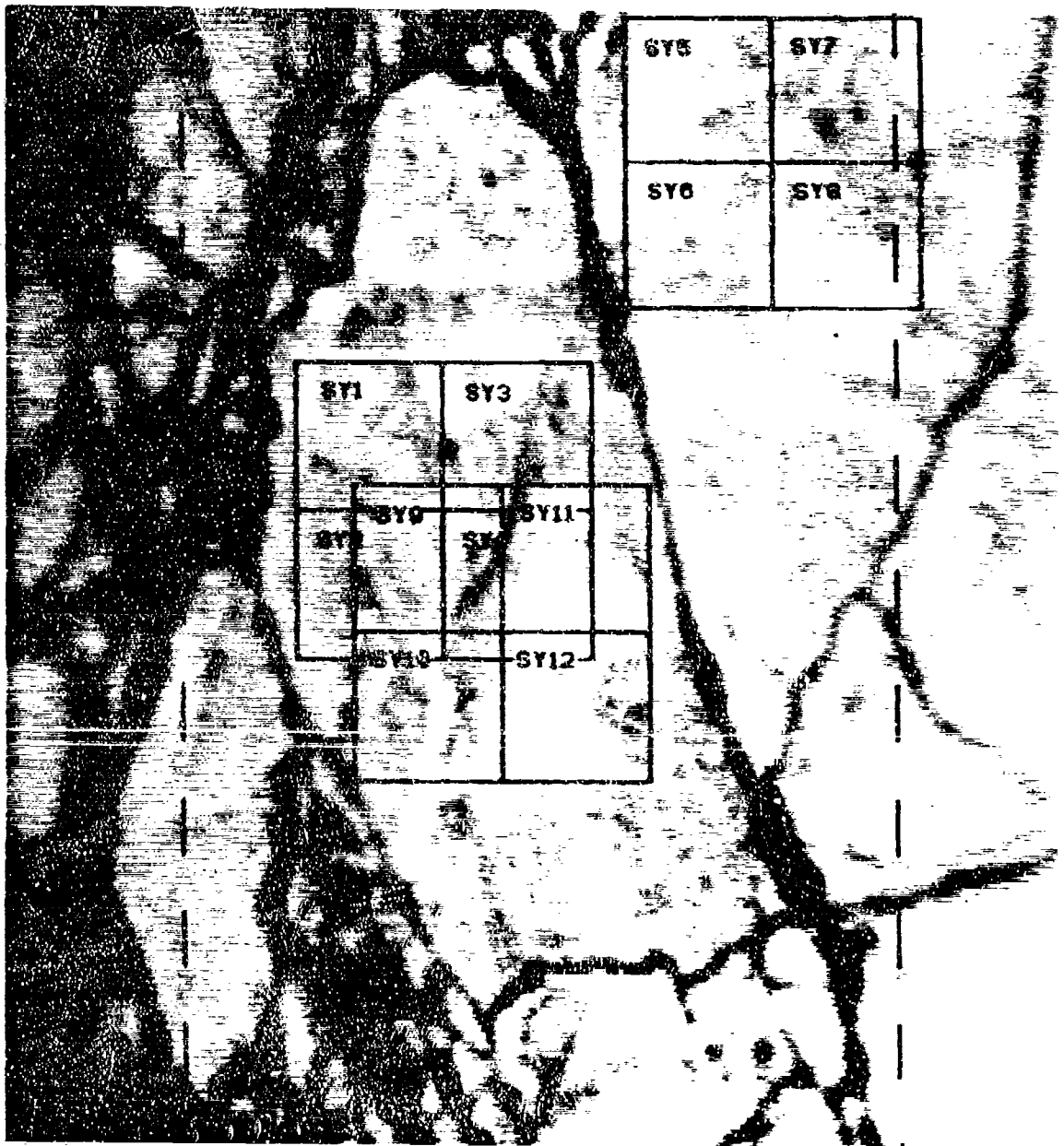


FIGURE IV-4. Full 512 x 512 pixel image showing locations where 64 x 64 pixel subareas were extracted as samples SY1 through SY12 of data set A.



FIGURE IV-5. Full 512 x 512 pixel image showing locations where 64 x 64 pixel subareas were extracted as samples FY1 through FY4 of data set A.



FIGURE IV--6. Full 512 x 512 pixel image showing locations where 64 x 64 pixel subareas were extracted as samples MY13 through MY18 of data set A.

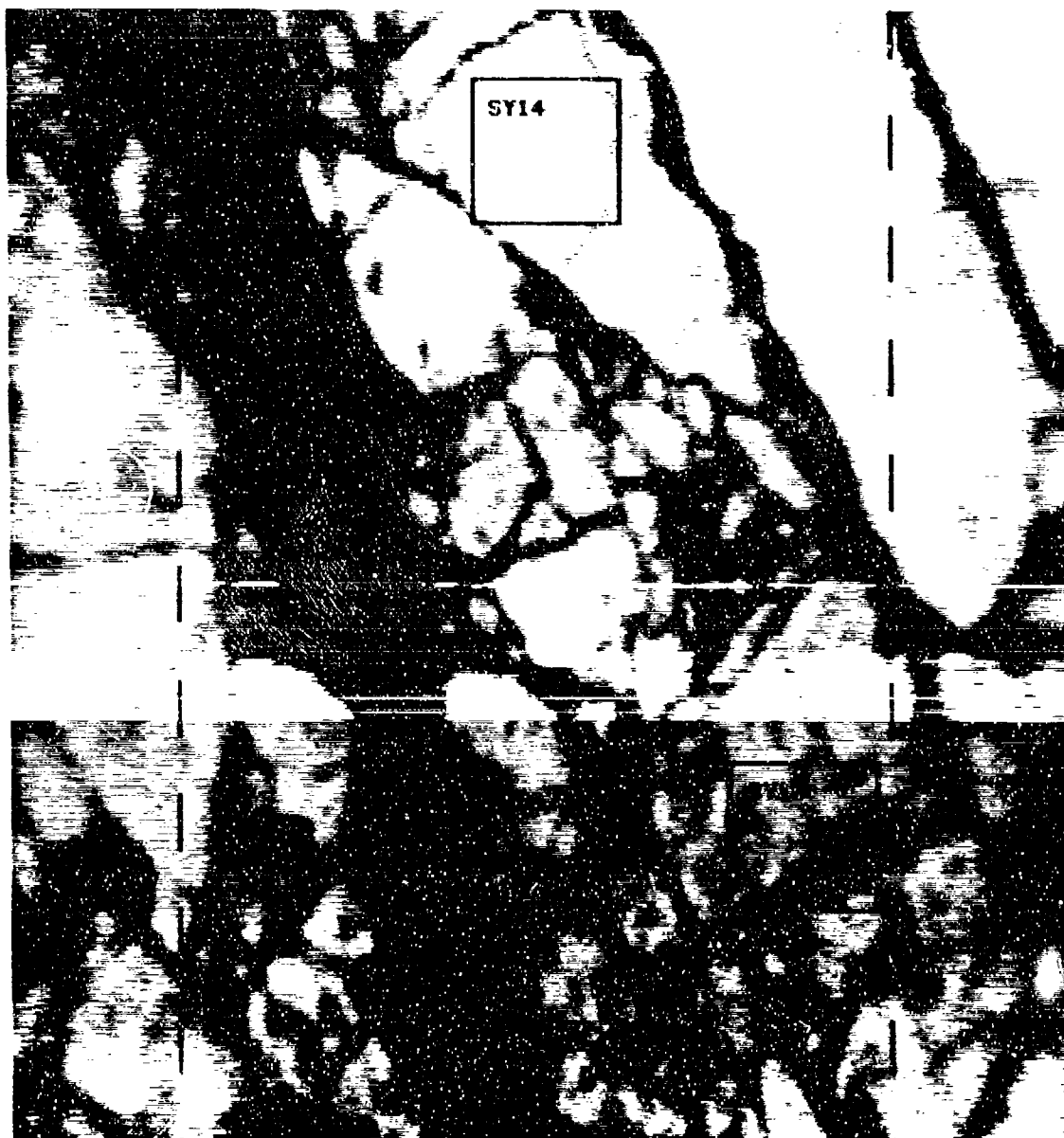


FIGURE IV-7. Full 512 x 512 pixel image showing locations where 64 x 64 pixel subareas were extracted as samples SY13 and SY14 of data set A.



FIGURE IV-8. Full 512 x 512 pixel image showing locations where 64 x 64 pixel subareas were extracted as samples SY15 through SY18 of data set A.



FIGURE IV-9. Full 512 x 512 pixel image showing locations where 64 x 64 pixel subareas were extracted as samples SY19 through SY22 of data set A.



FIGURE IV-10. Full 512 x 512 pixel image showing locations where 64 x 64 pixel subareas were extracted as samples SY23 through SY25 of data set A

from which the 50 subimages of data set A were extracted. The figures also contain graphic inlays marking the exact location of each of the extracted images and its associated assigned sample identification number. First-year ice samples are labeled FY1 through FY7, second-year samples are designated SY1 through SY25, and multiyear samples bear labels MY1 through MY18. Figures IV-1 through IV-10 also contain vertical dashed lines. These vertical lines mark the 100th sample in from each edge of the image. In Section II, considerations of image distortion and resolution lead to the recommendation that the outer 100 samples be ignored in the analysis. To a large degree this recommendation has been followed. However, the reader will note that some of the extracted samples overlapped into the forbidden edge areas. The violation of the guideline was unavoidable in these cases, but the number and amount of the overlaps into the edge regions is small and should not significantly impact upon the analysis results.

First-year ice is very uniform, so seven samples are considered adequate to characterize this ice type for the present study. Variability within ice categories increases with ice age, which leads to larger sampling requirements to reliably characterize the older ice types. Only 18 subimages of multiyear ice could be found for inclusion in data set A. Unlimited samples of second-year ice are available, but it makes no sense to obtain extremely large numbers of second-year samples when multiyear samples are limited to 18. Therefore, 25 second-year samples were selected to represent that ice type. Further, not all samples are unique. Several sample pairs overlap to some extent, producing samples with part of their area in common. In many cases, four samples were selected to be adjacent in the original image so that these adjacent 64×64 pixel samples could be treated as a single 128×128 pixel sample. The existence of both 64×64

and 128×128 images in data set A permits some limited investigation into the effects of sample size. Appendix C contains numeric values for each of the 25 candidate texture measures listed in Table III-1 for each of the 50 sample images in data set A.

B. NORMALITY

The topic of multivariate normality of the image texture data set was discussed in Section III.B.2. That section identified the linear correlation coefficient of the Q-Q plot as the chosen measure of normality. Q-Q plots, like the example shown in Appendix B, were constructed for each of the 25 variables in data set A. Each variable resulted in two plots, one for second-year ice and one for multiyear ice. (Q-Q plots were not constructed for first-year ice because of the small sample size.) A histogram of the linear correlation coefficients of the 50 Q-Q plots from data set A is shown in Figure IV-11. The majority of the values fall in the range of 0.95 to 1.0. It will be assumed that these are the normally distributed variables. Eight Q-Q plots have correlation coefficients below 0.95. These plots require further examination.

Four of the cases with correlation of 0.91 or 0.92 are SY-COR15, SY-LM1, MY-ENG15, and SY-KURT. For COR15, LM1, and KURT the multiyear correlation coefficient values are 0.97, 0.98, and 0.95, respectively, and the second-year coefficient for ENG15 is 0.98. Because the low correlation did not occur for both the second-year and multiyear cases, it is concluded that these low values result from the small sample sizes and are not truly indicative of nonnormality of these variables. However, the remaining four cases with Q-Q plot linear correlation coefficients in the 0.88 to 0.91 range are PR3 (both second-year and multiyear) and PR15 (both second-year and multiyear). Because all cluster prominence

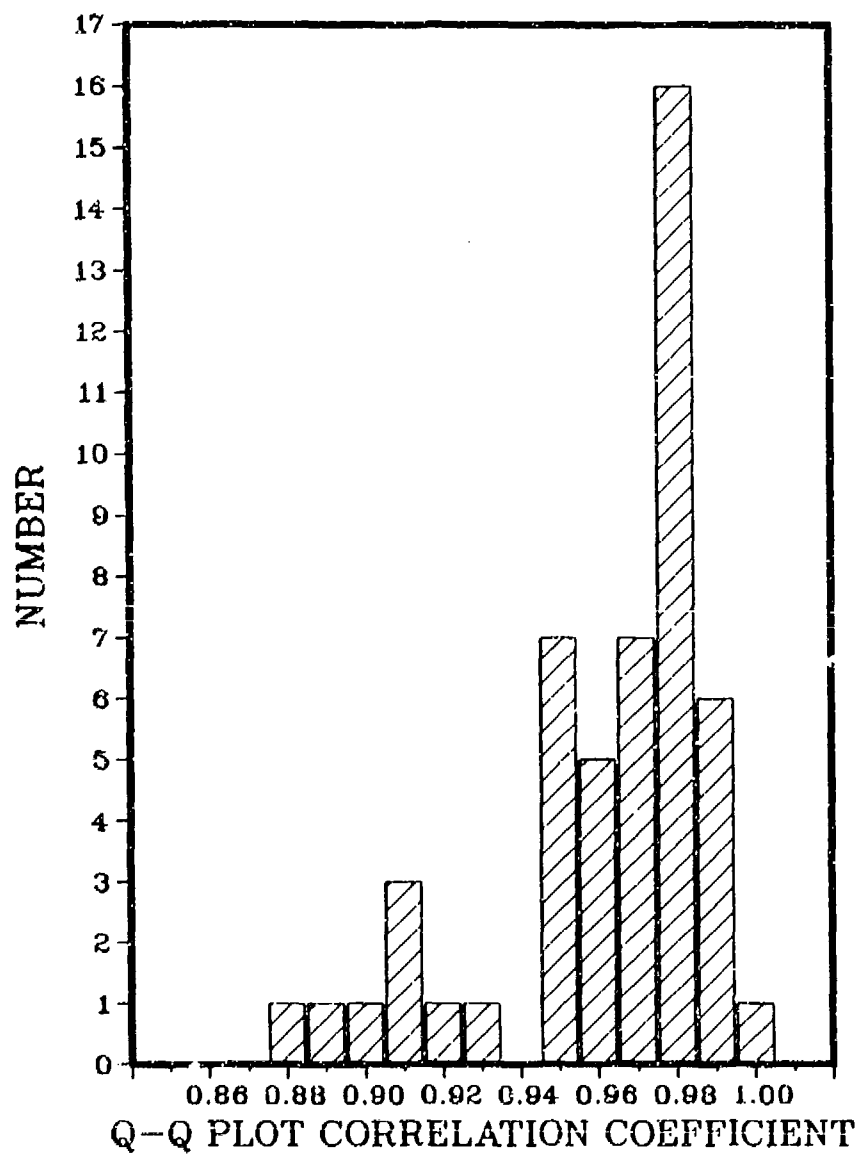


FIGURE IV-11. Histogram of linear correlation coefficient values for Q-Q plots of data from data set A.

cases show low correlation, it is concluded that cluster prominence does have a definite nonnormality problem, which should be corrected.

Examination of the histograms of the cluster prominence values shows long tails in the positive direction. This type of nonnormality can often be corrected by a logarithmic transformation of the original variables. Application of the equations

$$LPR3 = \ln(PR3 \times 10^{-6}) + 5$$

and

(IV - 1)

$$LPR15 = \ln(PR15 \times 10^{-6}) + 5$$

resulted in two new variables, LPR3 and LPR15, from the original PR3 and PR15 values. Q-Q plots of LPR3 and LPR15 resulted in four correlation coefficient values ranging from 0.95 to 1.0, which is in the range that has been assumed to represent normality. Therefore, the LPR3 and LPR15 variables have been added to Table C-1 in Appendix C. The new variables will replace PR3 and PR15 in all subsequent analysis. Data set A, after this transformation of cluster prominence values, is considered to be normally distributed within the limits of our ability to measure normality of data sets of this relatively small size.

C. CORRELATION WITH ICE TYPE

Correlation between ice type and the numerical values of texture measures was included in the list of desirable characteristics for texture measures in Section III-B. This ability of texture measures to serve as discriminators of ice type is, of course, the central requirement of this research area. The Bayesian ice-type classification based on each individual texture measure was chosen as the measure of correlation between ice type and the candidate texture measures.

Appendix A contains the theory of maximum Bayesian classification. Now that the approximate normality of data set A has been established in the preceding section (Bayesian classification assumes normally distributed variables), the ice-type classification task can begin. The hard clustering of data set A will be examined as a means of estimating the correlation of texture measure values with ice type. Hard clustering and Bayesian results will be compared. The clustering approach is introduced because it can be extended more easily to higher dimensional feature spaces.

The Bayesian classifier described in Appendix A represents a parametric approach to estimating classification accuracy. The class separation thresholds are calculated from two parameters: the means and standard deviations for each ice type for a given texture measure (see Eqs. (A-5) and (A-6)). Therefore these statistics have been calculated by ice type for each variable in data set A. The statistics are given in Table IV-1. Based on these statistics, Bayesian interclass thresholds are calculated using Eqs. (A-5) and (A-6). The thresholds become the limits of integration in Eq. (A-7), which results in the elements of the confusion matrix from which the classification accuracy is calculated according to Eq. (A-10). Table IV-1 shows these calculated ice-type classification accuracies for each of the 25 candidate texture measures.

Note that MEAN is the best ice-classification variable with an estimated accuracy of 93.0%. However, MEAN is not a true texture measure in the strict sense of the word. Of those texture measures other than MEAN, ENG3 and ENT3 scored the highest accuracies of 83.5% and 83.3%, respectively. Five of the texture variables have no calculated Bayesian classification accuracies given in Table IV-1. These five variables are all poor classifiers of ice types whose distributions overlap so strongly that Eq. (A-6) for classification threshold did

TABLE IV-1. DATA SET A STATISTICS

Meas	FY		SY		MY		Bayesian Accuracy	Clustering Accuracy
	Mean	Std	Mean	Std	Mean	Std		
MEAN	0.433	0.0411	1.689	0.1159	1.295	0.2049	0.931	0.88
VAR	0.379	0.0488	0.854	0.3826	1.631	0.5670	0.806	0.76
SKEW	3.120	0.1344	2.698	0.3715	2.448	0.3155	0.649	0.60
KURT	2.996	0.1723	3.421	0.5751	3.536	0.8060	-	0.50
INR3	0.451	0.0220	0.589	0.1169	0.895	0.2049	0.815	0.74
INR15	0.632	0.0789	1.473	0.6171	2.775	0.9244	0.820	0.78
LPR3	1.596	0.3572	3.444	1.0518	4.932	0.8165	0.806	0.74
LPR15	1.203	0.2970	2.598	0.9635	3.914	0.7502	0.781	0.70
HOM3	1.673	0.0732	1.537	0.1099	1.321	0.1556	0.717	0.66
HOM15	1.450	0.0819	1.059	0.1876	0.785	0.1523	0.817	0.78
ENG3	2.591	0.2109	1.762	0.3944	1.128	0.2731	0.835	0.78
ENG15	2.528	0.2615	1.569	0.4308	0.979	0.2560	0.823	0.76
ENT3	6.214	0.0817	6.626	0.2288	7.043	0.2099	0.833	0.80
ENT15	6.227	0.1042	6.729	0.2690	7.152	0.2155	0.821	0.78
COR3	4.008	0.6032	6.140	1.0932	7.127	0.5120	0.774	0.66
COR15	1.754	0.5246	1.337	1.2326	1.354	0.8284	-	0.50
LM1	0.129	0.0334	0.531	0.4162	1.000	0.4900	0.724	0.68
LM2	0.097	0.0202	0.335	0.1734	0.644	0.2490	0.801	0.72
LM3	0.395	0.0563	0.652	0.1925	0.986	0.3595	0.751	0.68
LM4	0.691	0.0968	0.685	0.1077	0.888	0.2254	-	0.60
LM5	2.668	0.2997	2.207	0.2866	2.548	0.3861	-	0.46
LM6	1.875	0.1848	1.534	0.2150	1.678	0.2925	-	0.40
LM7	2.259	0.2191	1.906	0.2999	1.958	0.3520	-	0.42
EDEN	1.841	0.1044	2.067	0.2072	2.519	0.2625	0.747	0.62
MBRL	3.475	0.1540	4.695	0.6889	5.794	0.7692	0.812	0.80

not have any useful solutions. ("useful solutions" means threshold values that lie between the means of the two distributions.)

An alternative to the parametric Bayesian calculations would be to use clustering techniques to actually classify data set A and compare the classification results with the known ice types of each sample. This alternative was pursued here because the clustering approach can be extended to give classifications in higher dimensional feature spaces more easily than can the Bayesian approach. The clustering scheme chosen is a hard clustering with normalized distance, d_i , as the cluster membership criterion.

$$d_i = \frac{|x_n - \mu_i|}{\sigma_i}, \quad (IV - 2)$$

where x_n is the value of the n^{th} data point being assigned class membership, μ_i and σ_i are the *a priori* mean and standard deviation of cluster i . Sample n will be assigned to that cluster for which $d_i < d_j$ for all $j \neq i$. This clustering technique should give results identical to the Bayesian values if (a) the variables are normally distributed, (b) the data set is large, and (c) all classes are represented equally in the ensemble.

The statistics from Table IV-1 have been used to supply μ_i and σ_i values for Eq. (IV-2), and all samples have been clustered according to the resulting distance values. The accuracies of the class assignments made by clustering the 50 samples of data set A are included as the last column in Table IV-1.

Comparison of the classification accuracies derived by the Bayesian and clustering methods shows that the Bayesian results are consistently higher than the empirical estimates based on clustering of data set A. The reason for this difference (hypothesized but not proven here) is that the Bayesian calculations

were based on the assumption of equal *a priori* probabilities for each of the possible classes. In data set A the first-year class is represented by 7 samples as compared to 25 and 18 samples, respectively, for the second-year and multiyear classes. The reason that first-year ice was sparsely represented is that it is easily distinguishable from the other two, which means it can be easily classified as first-year. Thus, the data set upon which the clustering was performed contained an underrepresentation of the easily classified samples relative to the more difficult samples to classify. This imbalance in ice type representation leads to an underestimate of approximately 6% in the classification accuracy.

The computational requirements for Bayesian classification become prohibitive for high dimensional feature spaces. Therefore, the clustering approach is used to examine how classification accuracy increases as texture measures are applied in combination. Clustering can be extended to higher dimensions by simply replacing x_n and μ_i with vectors \vec{X}_n and $\vec{\mu}_i$ and defining the distance metric as

$$d_i = (\vec{X}_n - \vec{\mu}_i)\Sigma^{-1}(\vec{X}_n - \vec{\mu}_i)^T, \quad (IV - 3)$$

where Σ is the covariance matrix of \vec{X}_n . Eq. (IV-3) has been applied to all pairwise combinations of the 25 texture measures calculated for data set A. These bivariate classification accuracy values are given in Table IV-2. The interesting thing to notice in this table is the manner in which the classification accuracies changed in comparison to the single variable cases. Several variable pairs involving the MEAN variable gave 98% accuracy compared to 88% for MEAN and 80% for the best of the true texture measures taken individually. Combining texture with mean intensity resulted in significant improvement. However, if we exclude MEAN and consider only the remaining variables, the MBRL/LM5 combination gave the best classification accuracy at 86%. All combinations with ac-

curacy values of 84% or higher are given in boldface numerals in Table IV-2. All of these better combinations (ENG3/HOM15, ENT3/COR15, MBRL/HOM3, MBRL/LM5, MBRL/LM6, and SKEW/LM2) involve one good univariate classification variable with one very poor univariate classifier. It is noteworthy that two good variables do not combine to form the best bivariate combinations. This fact leads to two conclusions. First, the better single variables are, in essence, alternate representations of the same texture information. Therefore, combining these variables introduces no new information for ice-type classification. The second conclusion is that it would be wrong to discard some of the very poor classification variables (a step which the author had intended to perform) because in the later multivariate analysis, the worst texture measures in the univariate sense may well become important.

Once the multivariate clustering software was in place for the bivariate calculations, it was a simple extension to combine all 25 texture measures into the clustering problem. By using all 25 of the variables of data set A, ice types could be classified with 100% accuracy. The MEAN variable was removed from the calculations and a 24-dimensional clustering of the true texture variables was performed. This 24-dimensional clustering resulted in a 94% classification of data set A. A word of caution is in order when considering the adequacy of our set of texture measures. First, remember that the image sample size is 64×64 pixels. The evaluation of classification accuracies is inexorably tied to sample size. As sample size increases, noise in the texture data is reduced and the clusters tighten, thereby improving interclass separability. However, this accuracy evaluation is conducted on the same data set that was used to design the classifier. An independent data set for accuracy evaluation would give more reliable and probably lower estimates than the procedure followed

here. Furthermore, the present data are from one location and one day. If data from different days and from different geographical regions were included, then numerical values for clustering accuracy in the ensemble of texture measures used here would undoubtedly change. Therefore, the reader is cautioned against placing too much importance on the exact values of clustering accuracies. The accuracies are presented here only to give a rough idea of the usefulness of various variables or combinations of variables. Conclusions beyond this are not supported by the data or the analysis procedures.

D. LINEAR MIXING

The global approach to image texture analysis developed here requires that texture measures mix linearly. This requirement was discussed in Section III.B.3. To evaluate the mixing characteristics of the candidate texture measures, eight samples from data set A are used. Samples FY1 through FY4 were combined into a single image, and samples MY13 through MY16 are combined into another image. These sample combinations resulted in two 128×128 pixel images with widely different textures, which were mosaicked to form the "mixture" image shown in Figure IV-12. Values of the candidate texture measures were computed for this synthetic ice-type mixture, as well as for the individual pure ice-type halves of the image. These results are given in Table IV-3. From these three measured values the percentage of the mixture that was multiyear ice (assuming linearity) was calculated and presented in the last column of Table IV-3. Of course, the correct answer for percentage of multiyear ice in the mixture is 50%. The texture variables, which resulted in approximately 50% multiyear compositions based on linear predictions, have their values shown in boldface type in the last column of the table.

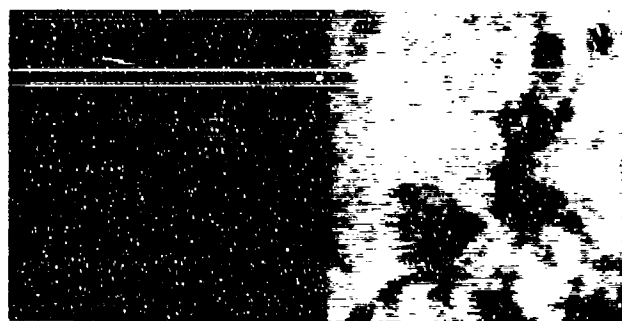
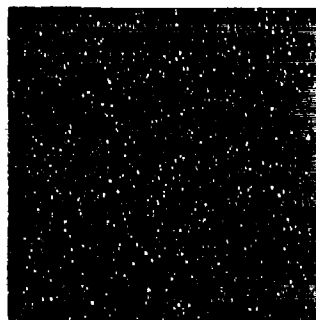


FIGURE IV-12. A FY (left) / MY (right) mixture image. The difference in brightness levels between the two halves prevents enhancing both halves simultaneously. Top image shows FY ice, bottom image shows MY ice.

TABLE IV-3. TEXTURE MEASURES FROM FIG. IV-12

Meas	S.F.	Measured Values			Linearly Calc.† MY Fraction
		FY Half	MY Half	Whole Image	
MEAN	10 ²	0.448	1.561	1.005	0.50
VAR	10 ²	0.424	1.499	31.954	29.30
SKEW	10 ⁰	3.061	2.200	3.046	0.02
KURT	10 ⁰	2.859	4.086	1.104	-1.43
INR3	10 ²	0.459	0.777	1.257	2.50
INR15	10 ²	0.657	2.430	5.049	2.48
LPR3	10 ⁰	1.858	5.066	10.164	2.59
LPR15	10 ⁰	1.463	4.174	10.098	3.19
HOM3	10 ⁻¹	1.651	1.404	1.518	0.54
HOM15	10 ⁻¹	1.417	0.858	1.104	0.56
ENG3	10 ⁻³	2.140	1.025	0.780	1.22
ENG15	10 ⁻³	2.015	0.750	0.650	1.08
ENT3	10 ⁰	6.412	7.223	7.531	1.38
ENT15	10 ⁰	6.476	7.527	7.768	1.23
COR3	10 ⁻¹	4.573	7.409	9.805	1.84
COR15	10 ⁻¹	2.161	2.038	9.244	-57.59
LM1	10 ²⁰	0.168	1.269	36.684	33.17
LM2	10 ¹⁶	0.103	0.557	0.320	0.48
LM3	10 ¹³	0.367	1.144	0.751	0.49
LM4	10 ¹³	0.626	0.989	0.814	0.52
LM5	10 ⁹	2.465	2.670	2.575	0.54
LM6	10 ⁶	1.785	1.751	1.753	0.94
LM7	10 ⁷	2.181	2.126	2.132	0.89
EDEN	10 ¹	1.847	2.303	2.302	1.00
MBRL	10 ⁰	3.731	6.222	25.525	8.75

† Linearly calc. MY fraction of mixture image is given by $(col5 - col3)/(col4 - col3)$.

Note in Table IV-3 that only the MEAN and 6 texture measures (HOM3, HOM15, LM2, LM3, LM4, and LM5) come close to linear behavior in the mixing operation. This fact is significant in light of intentions to apply linear unmixing theory in a global approach to analysis of image texture. The development of new image texture measures that possess linear mixing properties should be pursued if the global approach to texture analysis is to receive broad application.

The nonlinear behavior exhibited by the data in Table IV-3 was discovered to result mainly from the seam between the two halves of the mixture image. The large step in intensity arising from different mean intensities for the image halves completely dominates the value of many of the texture measures, even though the pixels bordering the seam occupy less than 1% of the image area. To verify that the nonlinearity results from the seam, intensities from both halves of the mixture image were offset to have a mean value of 128 digital counts. This new mixture image, with matched mean intensities, is shown as Figure IV-13. Texture measure calculations were repeated on this matched mean image with the results shown in Table IV-4. With the background intensity step removed, the number of texture measures that mix linearly is increased to 12.

Knowing that mean removal results in an increased number of usable texture measures, a decision must be made. One can either proceed with a greatly reduced number of texture variables, or produce an acceptable technique for detrending mixture images. It is not desirable to restrict the variety of texture measures under analysis, but detrending natural ice-type mixtures in KRMS images is a major area of study in itself. The approach chosen here was to use the largest possible number of variables while working with synthetic mixtures, since mean removal in synthetic mixtures is a trivial problem. It will then be demonstrated, using synthetic mixtures, that the smaller set of naturally mixing

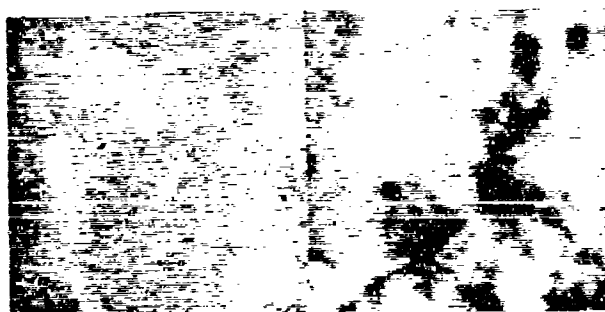


FIGURE IV-13. A FY (left) / MY (right) mixture image with the mean brightness of each half offset to a value of 128.

TABLE IV-4. TEXTURE MEASURES FROM FIG. IV-13

Meas	S.F.	Measured Values			Linearly Calc.† MY Fraction
		FY Half	MY Half	Whole Image	
MEAN	10 ²	1.278	1.281	1.280	..
VAR	10 ²	0.424	1.499	0.962	0.50
SKEW	10 ⁰	3.061	2.200	2.260	0.93
KURT	10 ⁰	2.859	4.086	5.186	1.90
INR3	10 ²	0.459	0.777	0.626	0.53
INR15	10 ²	0.657	2.430	1.528	0.49
LPR3	10 ⁰	1.858	5.066	4.395	0.79
LPR15	10 ⁰	1.463	4.174	3.521	0.76
HOM3	10 ⁻¹	1.651	1.404	1.522	0.52
HOM15	10 ⁻¹	1.417	0.858	1.136	0.50
ENG3	10 ⁻³	2.140	1.025	1.390	0.67
ENG15	10 ⁻³	2.015	0.750	1.155	0.68
ENT3	10 ⁰	6.412	7.223	6.960	0.68
ENT15	10 ⁰	6.476	7.527	7.178	0.67
COR3	10 ⁻¹	4.573	7.409	6.741	0.76
COR15	10 ⁻¹	2.161	2.038	2.011	1.22
LM1	10 ²⁶	0.168	1.269	0.679	0.48
LM2	10 ¹⁶	0.103	0.557	0.320	0.48
LM3	10 ¹³	0.368	1.145	0.752	0.49
LM4	10 ¹³	0.626	0.989	0.814	0.52
LM5	10 ⁹	2.466	2.670	2.576	0.54
LM6	10 ⁸	1.785	1.751	1.753	0.94
LM7	10 ⁷	2.181	2.126	2.132	0.89
EDEN	10 ¹	1.847	2.303	2.088	0.53
MBRL	10 ⁰	3.731	6.222	4.807	0.43

† Linearly calc. MY fraction of mixture image is given by $(col5 - col3)/(col4 - col3)$.

texture measures does not significantly reduce unmixing accuracy in this data set. We will therefore proceed to natural mixtures with a reduced variable set and avoid dealing with the problem of detrending natural KRMS mixture images.

E. INVARIANCE UNDER LINEAR TRANSFORMATION

Section III.B.5 discussed invariance under linear transformation of image intensity values as a desirable but nonessential characteristic of image texture measures. A 128×128 pixel image formed from data set A samples MY13 through MY16 was used to investigate this aspect of the candidate texture measures. Numerical values of the texture measures for this image were given in Table IV-3. Intensity values within this image were then linearly transformed according to Eq. (III-3), where the constants a and b are given values of 0.8 and -20.0, respectively. Texture measures were recalculated for the transformed image. Table IV-5 compares the texture measures calculated from the original and from the transformed imagery, as well as gives a percentage difference between the two values. From these difference values, only SKEW, KURT, COR3, and COR15 are clearly invariant under linear transformation of intensities. Three additional variables, ENT3, ENT15, and MBRL, are probably close enough to invariant to be considered as such in the present application. However, these invariant texture measures, with the exception of MBRL, are not those that mix linearly. Therefore, numerical values of texture measures or coefficients involved in ice-type discrimination algorithms based on texture in this KRMS data set are not universal. The philosophy of texture analysis developed in this dissertation will be applicable to other data sets, but the specific numerical values derived from this work will need to be recalculated for other data sets.

The development of image texture measures that are invariant to linear

TABLE IV-5. EFFECT OF INTENSITY TRANSFORMATION ON
IMAGE TEXTURE MEASURES

Meas	S.F.	Calculated Values		Difference (%)
		Orig. Image	Trans. Image	
MEAN	10^2	1.561	1.049	-32.8
VAR	10^2	1.499	0.960	-36.0
SKEW	10^0	2.200	2.201	0.0
KURT	10^0	4.086	4.079	-0.2
INR3	10^2	0.777	0.499	-35.8
INR15	10^2	2.430	1.556	-36.0
LPR3	10^0	5.066	4.172	-17.6
LPR15	10^0	4.174	3.282	-21.4
HOM3	10^{-1}	1.404	1.744	24.2
HOM15	10^{-1}	0.858	1.088	26.8
ENG3	10^{-3}	1.025	1.935	88.8
ENG15	10^{-3}	0.750	1.400	86.7
ENT3	10^0	7.223	6.708	-7.1
ENT15	10^0	7.527	7.028	-6.7
COR3	10^{-1}	7.409	7.402	-0.1
COR15	10^{-1}	2.038	2.037	0.0
LM1	10^{20}	1.269	0.812	-36.0
LM2	10^{16}	0.557	0.356	-36.1
LM3	10^{13}	1.144	0.731	-36.1
LM4	10^{13}	0.989	0.632	-36.1
LM5	10^9	2.670	1.710	-36.0
LM6	10^8	1.751	1.129	-35.5
LM7	10^7	2.126	1.380	-35.1
EDEN	10^1	2.303	1.845	-19.9
MBRL	10^0	6.222	6.575	5.7

transformation of image intensity is highly desirable and is recommended as a subject for further study.

F. SUMMARY OF THE TEXTURE ANALYSIS DATA SET

A set of fifty 64×64 pixel KRMS images, each containing only one of three possible ice types (FY, SY, and MY), has been defined and named data set A. The 25 candidate texture measures discussed in Section III were calculated for each of these samples. The numerical values of these texture measures were analyzed to lead to the following conclusions.

- All texture measures are approximately normally distributed. (The cluster prominence variables required logarithmic transformation to achieve normality.)
- Most of the texture measures do not mix linearly. A mean removal step was shown to solve the nonlinearity problem for some variables. Texture measures with linear mixing properties, without mean removal, are recommended as an area of future study.
- Texture measures, as a general rule, are not invariant to linear transformations of image intensity values. This means that KRMS calibration, as well as day-to-day environmental variability, cannot be ignored. Universal algorithms are not to be expected from the present ensemble of image texture parameters. A "recalibration" of image texture analysis algorithms is indicated for each new KRMS data set.
- The 25 candidate texture measures proposed in Section III do seem adequate in the sense that they contain sufficient texture information to reasonably identify the ice type of the samples in data set A. In other words, successful

clustering of data set A indicates that our texture measures form the basis of an effective feature space, as "effective" is illustrated in Figure III-2.

- Very little improvement in ice-type discrimination power is observed as univariate clustering progresses to multivariate clustering of the texture feature vectors. This lack of improvement indicates that the various texture measures are highly correlated with each other or, in other words, many texture measures indicate the same type of texture information. The true dimensionality of the data set is therefore seen to be significantly less than the number of texture measures included in the study. This correlation between variables indicates that farther along in the study, a large reduction in dimensionality is expected from factor analysis, and that a relatively small number of end members will be required to model the data set as an ensemble of linear mixtures of end members.

Of the 25 texture variables calculated for data set A, only 12 will be carried on into the analysis of synthetic mixtures that will have means removed. The 13 variables eliminated are the result of failure to mix linearly. As the study progresses to natural mixtures, the 12 variables will be further reduced to 7. This result emphasizes the need for further investigation into new texture measures with linear mixing properties.

The one characteristic of texture measures identified as desirable in Section III, but not addressed in this Section, is the constant sum requirement imposed on texture feature vectors by the CABFAC and QMODEL computer codes, which are utilized for factor analysis and linear unmixing in later analysis steps. The constant sum problem is the subject of Section V.

V. THE CONSTANT-SUM CONSTRAINT

The texture feature vectors representing each sample image in data set A have been developed for further analysis using the linear unmixing model contained within the CABFAC and QMODEL computer codes. (CABFAC and QMODEL processing is not encountered until Section VI.) However, the constant-sum constraint, which is an integral part of the CABFAC/QMODEL implementation of linear unmixing, will be the topic of this section. Data set A is not constant-sum. Summations of the columns of Table C-1 range from 47.8 to 75.1, so this problem must be addressed.

The constant-sum constraint can be traced to Miesch (1976a, 1976b, 1976c), who approached the unmixing problem by associating axes of an oblique feature space with end members, and vector components in this end-member coordinate system with end-member compositions in each sample. However, linear unmixing can be performed on any linear data set; the constant sums are not a theoretical requirement. We will see how any general linear data set can be unmixed and how the constant-sum constraint arises from the Miesch approach.

Consider a multivariate data set consisting of sample vectors with M measurements comprising each vector. Sample vectors, X_i , consist of x_{i1}, \dots, x_{iM} , where x_{ij} represents the j^{th} measurement on the i^{th} sample. The constant-sum constraint can be written,

$$\sum_j x_{ij} = K, \quad (V - 1)$$

for all i , where K is a constant.

Linear mixing is illustrated by Figure (V-1a), where $M = 2$. If end members E_1 and E_2 are mixed in equal proportions to form sample X_1 , then

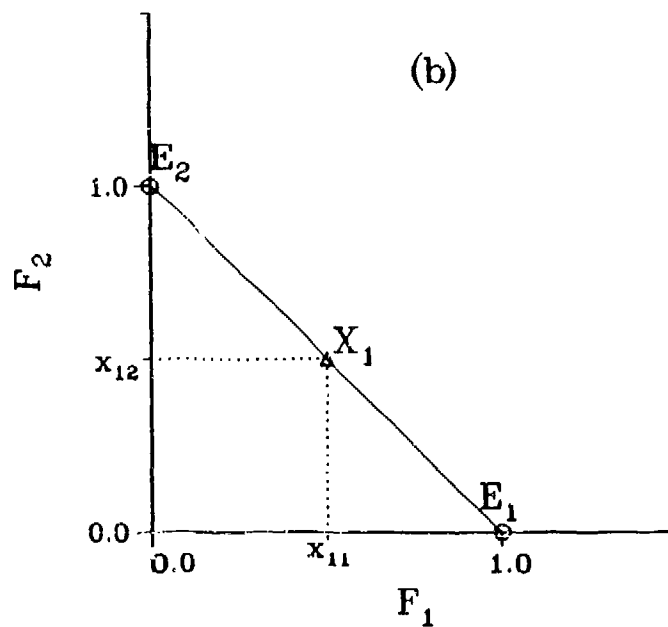
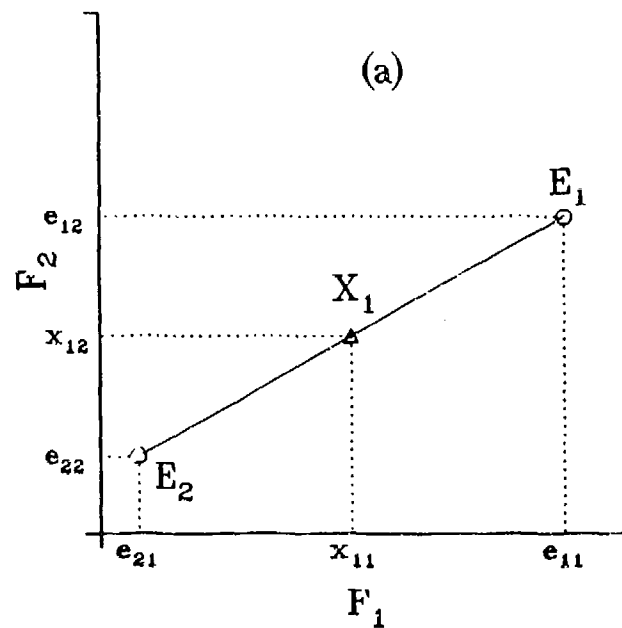


FIGURE V-1. One-dimensional linear mixing of (a) variable row sum data, and (b) constant row sum data. F_1 and F_2 are axes of the feature space. E_1 and E_2 are end members.

assuming individual measurements mix linearly,

$$\begin{aligned}x_{11} &= (e_{11} + e_{21})/2 \\x_{12} &= (e_{12} + e_{22})/2.\end{aligned}\tag{V - 2}$$

Clearly, if point $X_1 = (x_{11}, x_{12})$ is plotted in Figure (V-1a), it will fall equidistant from E_1 and E_2 on the line segment $\overline{E_1 E_2}$ that connects these end members. Any other samples that could possibly be formed by mixing E_1 and E_2 in varying proportions would similarly fall on the line $\overline{E_1 E_2}$. Thus, one gets physically meaningful unmixing results if the data set includes only points that fall on $\overline{E_1 E_2}$. If the samples do not lie in a line between the end members (or close to a line, allowing for noise in the data), then applying unmixing algorithms is unjustified because the data set is not made up of mixtures of E_1 and E_2 . The need for a linear constraint on the data set is, therefore, obvious. Equation (V-1), which expresses the constant-sum constraint, does indeed define a line. However, Eq. (V-1) represents only a subset of the possible lines that exist in a two-dimensional measurement space. Line $\overline{E_1 E_2}$ in Figure (V-1a) does not satisfy Eq. (V-1). Nonetheless, we have just shown that the linear equation (Eq. (V-2)) can be applied to form mixtures that fall along this line. The inverse procedure for unmixing data along $\overline{E_1 E_2}$ is equally straightforward. Therefore, $\overline{E_1 E_2}$ is clearly a valid mixing/unmixing line, even though it is not constant-sum. The only constraint on a data set for unmixing that arises from the mathematics is the general hyperplane equation

$$\sum_j a_j x_{ij} = K.\tag{V - 3}$$

The constant-sum constraint is a special case of Eq. (V-3), where $a_j = 1$ for all j . In other words, the constant-sum constraint is more restrictive than the mathematics of linear unmixing requires. Many data sets that satisfy Eq. (V-3)

and that could be unmixed in principle cannot be processed using the CABFAC/QMODEL programs because of the constant-sum requirements associated with that code.

If we require the end members to lie on the axes F_1 and F_2 at unit distance from the origin, as shown in Figure (V-1b), then the mixing line is also a constant sum line where $K = 1$. In this case the proportion of E_1 in sample X_1 is simply x_{11} and the proportion of E_2 in sample X_1 is x_{12} . The end members lie on the axes, and the coordinates of the samples are the end-member proportions, which is one advantage of constant sum data. In most cases the coordinate system will be oblique rather than orthogonal as shown in Figure (V-1b), but the same principle of associating sample coordinates with end-member compositions also applies in the oblique case. In the general case shown in Figure (V-1a) the calculation of end-member compositions is more complex. In the general unmixing case, end-member proportions are given by

$$\begin{aligned} E_1 \text{ proportion} &= (x_{11} - e_{21}) / (e_{11} - e_{21}) \\ E_2 \text{ proportion} &= (e_{11} - x_{11}) / (e_{11} - e_{21}). \end{aligned} \tag{V - 4}$$

The additional complexity associated with the general case can become significant as the dimensionality of the feature space becomes large and the basis vectors are oblique. The CABFAC/QMODEL programs were designed to take advantage of the simplification that results from constant-sum data matrices. In addition to the simplification just mentioned, other aspects of the CABFAC/QMODEL code developed over the years took advantage of the constant-sum assumption. For example, dimensionality reduction by identifying eigenvectors of the cosine θ matrix (discussed in Section VI) is only guaranteed to be a reasonable approach if the data are at least approximately constant-sum. Therefore, for several reasons, the constant-sum constraint cannot be violated when using these programs.

Klovan (1981) has published a generalization of Q-mode factor analysis for data matrices with variable row sums. The Klovan method involves projecting data points onto a plane of constant sum \bar{K} , where \bar{K} is the average row sum for the data matrix. K_i , the row sum for the i^{th} sample vector is saved for later use. The analysis proceeds in the \bar{K} plane. The resulting composition loadings and composition scores are then scaled by \bar{K}/K_i so that estimates of the raw data can be obtained in the original metric. The Klovan method would seem (to the author) to be useful only in those cases where the original data are very close to constant-sum and only small adjustments are required. For other cases where the data were in a hyperplane nearly perpendicular to the constant-sum plane, the Klovan method could result in a large loss of information from the data set. Therefore, this solution does not seem to be practical for the general variable row-sum problem. The Klovan method was not utilized in this study.

Full (1988) claims to have a general variable row-sum version of CABFAC/QMODEL operational at the time of this writing. However, this software was not available at the time that this analysis was performed. We were therefore forced to find alternatives to deal with the variable row-sum nature of this image texture data.

We have chosen to deal with the constant-sum problem by describing a data set transformation that would operate on linear data sets obeying, at least approximately, the general linear condition of Eq. (V-3) to produce constant-sum data that obey Eq. (V-1) precisely. Such preprocessed data would then be suitable for processing by CABFAC/QMODEL. The transformation, named the "Scale Invert then Force Transformation" (SIFT), will be demonstrated on a simple two-dimensional data set consisting of 10 samples. The SIFT will then be extended to hyperspace and applied to the usable variables from data set A to

form a new data set, called data set B, which will be passed into the CABFAC factor analysis in Section VI.

A. THE SCALE INVERT THEN FORCE TRANSFORM

1. The Concept

The mathematically rigorous approach to forming constant-sum data sets from linear data would be to find a best-fit hyperplane to the linear data set, and then find the rotation that would transform this best-fit plane to a constant-sum plane. If the original data were nearly linear, then the rotated data points would be close enough to constant-sum that they could be forced to true constant-sum without losing significant information content. The equation that forces a data set X_i to constant-sum set X'_i is given by

$$x'_{ij} = x_{ij} / \sum_j x_{ij}. \quad (V - 5)$$

To illustrate this "fit, rotate, then force" method, referred to hereafter as FRF, the hypothetical two-dimensional data set shown in Table V-1 will be analyzed. These data are mixtures of the end members $E_1 = (0.41, 0.40)$ and $E_2 = (1.25, 0.63)$, plus some noise so that the data set is not perfectly linear. This data set has no significance. These values were made up by the author in order to provide a simple data set of known end-member proportions, which could be used to illustrate the SIFT and FRF procedures and to show that these are similar, but not mathematically identical means for forming constant-sum data. End members E_1 and E_2 are samples 8 and 2, respectively, in Table V-1.

The fourth column of Table V-1 is the row sum for each of the samples. Note that the sums range from 0.81 to 1.88, clearly indicating the variable row-sum nature of this data. Column 5 of Table V-1 is the percentage of end-member

TABLE V-1. A HYPOTHETICAL TWO-DIMENSIONAL DATA SET

Sample	x_{i1}	x_{i2}	$x_{i1} + x_{i2}$	$\%E_1$ †
1	1.10	0.55	1.65	19.3
2	1.25	0.63	1.88	0.0
3	0.51	0.45	0.96	87.1
4	1.20	0.58	1.78	7.1
5	0.80	0.48	1.28	54.3
6	0.75	0.53	1.28	57.8
7	0.69	0.49	1.18	65.7
8	0.41	0.40	0.81	100.0
9	0.98	0.56	1.54	32.1
10	0.62	0.45	1.07	75.0

† $\%E_1$ is calculated by projecting the sample onto the line $\overline{E_1 E_2}$ and then applying Eq. (V-4).

E_1 that is present in each of the samples. Even though the hypothetical data set is not constant-sum, it is highly linear as evidenced by a linear correlation coefficient between x_{i1} and x_{i2} of 0.95. A least-squares linear fit to the data results in a line described by Eq. (V-3), where $a_1 = -0.72$, $a_2 = 3.12$, and $K = 1.0$. Figure V-2 is a plot of these data. The linear fit to the data and the "row sum = 1" line are also shown in the figure. From Figure V-2 we can see that a clockwise rotation of 58° would put the best-fit line into a constant-sum orientation. This two-dimensional rotation can be accomplished by subtracting the mean vector $\bar{X} = (0.831, 0.512)$ from each of the data vectors, premultiplying the mean-removed data vectors by the matrix

$$\begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix},$$

where θ is the angle of rotation, and then adding the mean vector to the result. The data set that results from this rotation about the mean to the constant-sum orientation, followed by a forcing to the "row sum = 1" line (Eq. (V-5)), is listed in Table V-2 and plotted in Figure V-3.

By comparing the last columns of Tables V-1 and V-2 we see that the FRF procedure has resulted in slightly altered proportions of end-member E_1 . The rms difference between E_1 compositions in Tables V-1 and V-2 is 2.2%. The rotation does not change the E_1 proportions, since a rigid rotation leaves relationships between samples unchanged; rather, the forcing to constant-sum via Eq. (V-5) alters compositional calculations. The rotation prior to application of Eq. (V-5) is necessary to minimize the errors in compositional estimates introduced by the forcing procedure. The need for the rotation is obvious if one pictures a linear data set laying along the $x_{i1} = x_{i2}$ line. In this case, forcing to constant-sum would collapse all data points to the same location on the constant-sum line,

TABLE V-2. DATA FROM TABLE V-1 AFTER THE FRF TRANSFORMATION

Sample	x'_{i1}	x'_{i2}	$x'_{i1} + x'_{i2}$	$\%E_1$ †
1	0.768	0.232	1.000	17.6
2	0.840	0.160	1.000	0.0
3	0.477	0.553	1.000	80.7
4	0.819	0.181	1.000	6.5
5	0.602	0.398	1.000	53.6
6	0.576	0.424	1.000	59.2
7	0.543	0.457	1.000	66.4
8	0.388	0.612	1.000	100.0
9	0.698	0.302	1.000	32.8
10	0.503	0.497	1.000	75.1

† $\%E_1$ is calculated by applying Eq. (V-4).

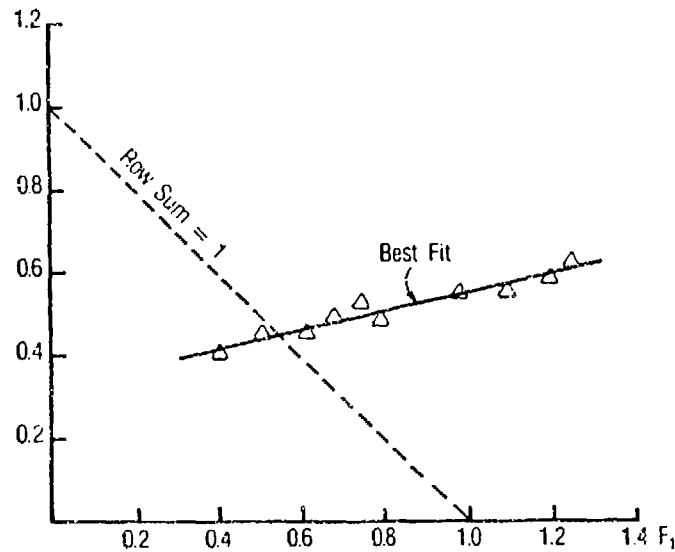


FIGURE V-2. Plot of the hypothetical data set listed in Table V-1. Least-squares linear fit to the data and constant-sum lines are also shown.

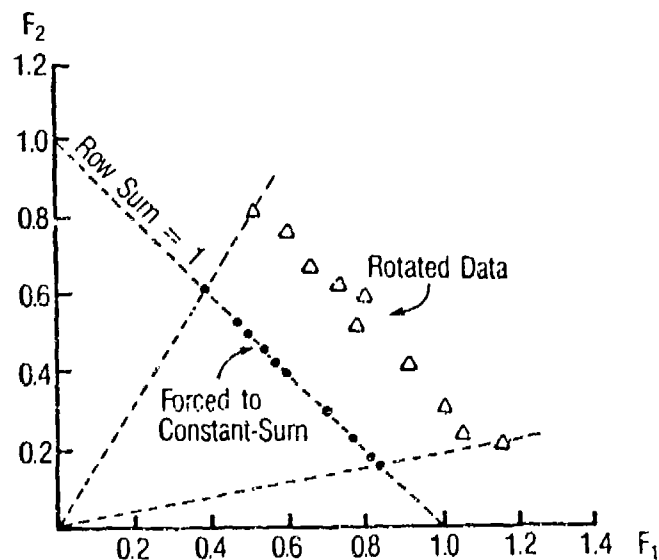


FIGURE V-3. Illustration of rotation and constant-sum forcing of the data shown in Figure V-2.

which would result in a total loss of compositional information contained in the original data.

The extension of the FRF approach to large data sets of high dimensionality is beset by mathematical complexity and computational burden. Therefore, an alternate approach to forming constant-sum data from general linear data sets is presented here. The proposed data transformation avoids multiple linear regression and hyperspace rotations, while giving nearly equivalent results, at least for the simple hypothetical data set examined here. The data set from Table V-1 will be used to show that useful constant-sum data can be obtained by three steps.

- Scale each variable over the range of 0 to 1.
- Invert x_{i2} by replacing x_{i2} with $(1 - x_{i2})$.
- Force data to constant sum of 1 using Eq. (V-5).

Figure V-4 graphically illustrates each step. Table V-3 contains the constant-sum data set thus derived. Comparison of the right column of Table V-3 with the same column in Table V-1 allows us to calculate the loss in mixing accuracy that results from this simple constant-sum transformation. The comparison shows that the percentage of E_1 in each sample in Table V-3 differs from the percentages in Table V-1 by an rms value of 5.5%. This value is compared to an error of 2.2% introduced by the more complex FRF transformation leading to the data in Table V-2. Thus, the performance of the SIFT is slightly inferior in this case to the more rigorous FRF algorithm. General conclusions concerning the relative accuracies of the two methods should not be drawn from this one example. The unmixing accuracy of the SIFTed data could have been forced to come out to any desired value by changing the values in the hypothetical data set accordingly.

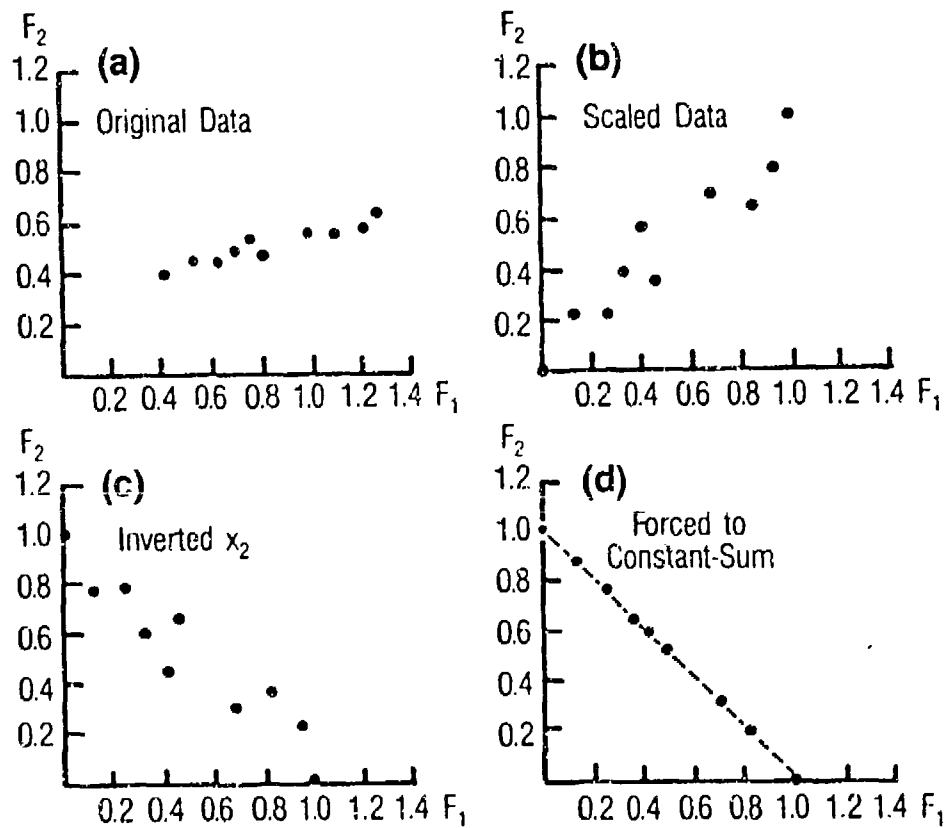


FIGURE V-4. Graphical representation of the scale (b), invert (c), and force (d) transformation (SIFT) of the data shown in Figure V-2.

TABLE V-3. DATA FROM TABLE V-1 WITH THE
SIFT APPLIED

Sample	x'_{i1}	x'_{i2}	$x'_{i1} + x'_{i2}$	$\%E_1$ †
1	0.702	0.298	1.000	29.8
2	1.000	0.000	1.000	0.0
3	0.132	0.868	1.000	86.8
4	0.812	0.188	1.000	18.8
5	0.416	0.584	1.000	58.4
6	0.482	0.518	1.000	51.8
7	0.354	0.646	1.000	64.6
8	0.000	1.000	1.000	100.0
9	0.691	0.309	1.000	30.9
10	0.242	0.758	1.000	75.8

† $\%E_1$ is calculated by applying Eq. (V-4).

The results of both the FRF transformation and the SIFT are data dependent. This comparison has been conducted only to demonstrate the SIFT procedure and to show that in some cases, such as the hypothetical data set used here, the SIFT and FRF results can be similar in terms of resulting unmixing accuracy. A multidimensional version of the SIFT will be applied to the image texture data analyzed in this study. The appropriateness of this algorithm for forming constant-sum data will tend to be further substantiated when we compare the unmixing results of raw data, forced to constant-sum, and SIFTed data in a later section. However, a precise estimation of the errors introduced by the SIFT of this image texture data can be established only by processing these data using a variable-sum version of CABFAC/QMODEL in the future.

The scaling of variables that is proposed in connection with the SIFT is not a new idea. The CABFAC code has contained an option to scale the data over the range 0 to 1 for many years (Klován and Imbrie, 1971). Scaling is normally performed to balance the weighting of variables when magnitudes are extremely different. The present texture variables range in value by many orders of magnitude, so a scaling of the data would be advisable even if the SIFT was not performed. The original contribution here is the inversion of one of the variables in order to bring the data set into a more nearly constant-sum orientation prior to forcing to constant-sum.

It should be reiterated that scaling and range inversion are both linear operations; therefore, the unmixing characteristics of a data set should not be changed in any way. Forcing to constant sum is a nonlinear process that can degrade unmixing calculations on the forced data set. Scaling and range inversion are performed to minimize the adverse effects that may result from forcing to constant-sum.

2. Extension to Higher Dimensions

The SIFT can be easily extended into hyperspace. For the M -dimensional case the scaling of all variables over the 0 to 1 range remains as the first step. However, the inversion step becomes more involved. To perform the inversion, one variable is selected as a reference variable. The choice of reference variable is arbitrary; for purposes of discussion we will assume the first variable is the reference variable. Correlation coefficients are then calculated between the reference variable and each of the $M - 1$ remaining variables. Certain variables are then inverted and others left unchanged in such a manner that half of the variables are positively correlated and the other half negatively correlated with the reference variable. (The $(1 - x_{ij})$ range inversion of a scaled variable will change the sign of its coefficient of correlation with respect to the reference variable.) The procedure used here is to make all odd-numbered variables positively correlated, and all even-numbered variables negatively correlated with the reference variable. The hypothesis put forth here is that if the original variable-sum data were nearly linear, then the scaled, half-inverted data will be close enough to constant sum that it can be forced the rest of the way with a minimal loss of information content.

3. Data Set B

The 12 texture measures identified in Section IV as being suitable for analysis of synthetic mixtures are extracted from data set A and SIFTed to form data set B, which will be the subject of the first unmixing studies. The row sums of data set B after range scaling, but prior to inversion or forcing, range from 2.40 to 8.55. VAR is selected as the reference variable. Table V-4 shows the correlation coefficients between VAR and the other 11 variables. INR3, HOM15, LM1, LM3,

**TABLE V-4. CORRELATION OF TEXTURE VARIABLES
WITH "VAR"**

Variable	r	Variable	r
VAR	1.000	LM2	0.843
INR3	0.937	LM3	0.734
INR15	0.962	LM4	0.572
HOM3	-0.779	LM5	0.122
HOM15	-0.877	EDEN	0.888
LM1	0.903	MBRL	0.878

LM5, and MBRL were range inverted to make odd- and even-numbered variables positively and negatively correlated with VAR. After inversion, row sums varied from 5.20 to 6.94, which is much closer to a constant value than were the row sums prior to inversion. The inverted data are then forced to a constant sum of 1, thereby completing the SIFT procedure. Appendix D contains the resulting data set B in tabular form. Two questions should be addressed in connection with the formation of data set B. Is the data set nearly linear so that application of the SIFT and the subsequent attempts at unmixing are justified? How much ice classification power has been lost by reducing the number of variables from 24 to 12?

The linearity of the data set has been inferred in Section IV, where it was observed that the ice classification accuracy based on clustering did not improve significantly as the dimensionality of the clustering feature space was increased. The linearity of the data is confirmed here by direct calculation. A fit of the 12 texture measures from data set B to the general linear hyperplane equation (Eq. (V-3)), by multiple linear regression, resulted in a multiple correlation coefficient of 0.991. This high value for multiple correlation coefficient confirms the general linear nature of data set B.

The loss of ice discrimination power resulting from the reduced number of texture of variables was investigated by clustering of the 12 variables from data set B, and comparing cluster memberships with known ice types for each sample. In Section IV, where similar clustering was performed using 24 of the data set A variables (MEAN was excluded), the resulting ice classification accuracy was 94%. Clustering of data set B resulted in a 92% classification accuracy, indicating that the loss of ice discrimination power in going from 24 to 12 variables was not significant.

Data set B is now ready for factor analysis in Section VI. Factor analysis will determine the true dimensionality of the data set and give an indication of the number of end members supported by the texture data.

VI. UNMIXING OF PURE SAMPLES

A. THE LINEAR MIXING MODEL

The linear mixing model can be stated,

$$\vec{X}_i \simeq \sum_{k=1}^K a_{ik} \vec{E}_k, \quad (VI-1)$$

where \vec{E}_k is the k^{th} end member. That is, sample vector \vec{X}_i can be approximated by a weighted sum of several (K to be exact) end-member vectors. A constraint on the weights, a_{ik} , is that

$$\sum_k a_{ik} = 1. \quad (VI-2)$$

If the \vec{E}_k vectors represent "pure" end members, then linear unmixing theory proceeds to interpret the a_{ik} coefficients as representing the proportions of end members \vec{E}_k present in sample i . This interpretation is consistent with the constraint of Eq. (VI-2), since the sum of the end-member proportions in a sample must be unity, i.e., 100% of every sample must consist of something. Since our stated objective is to apply unmixing theory to image texture data, the first question that arises is whether our data fit the unmixing model of Eq. (VI-1). The relevant questions are as follows.

- Eq. (VI-1) defines an approximation of \vec{X}_i by a weighted sum of end members. How close does this approximation need to be, to be useful?
- How many end members, i.e., what value of K is necessary to provide the desired accuracy in the approximation? In the trivial case where K is equal to the number of samples in the data set, Eq. (VI-1) reduces to an exact expression. However the exact solution, where every sample is an end member, is of no practical value. Unless the \vec{X}_i vectors can be approximated

with sufficient accuracy when K is relatively small, the unmixing analysis will usually be difficult to interpret and, therefore, will offer little new insight into the nature of the data set.

- What are the mixing proportions of the various \vec{E}_k s in each of the samples, i.e., what values of the a_{ik} coefficients will fit Eq. (VI-1) to our data set?
- Do the end members and mixing proportions of these end members in each sample lead to improved understanding of the physical nature of the data?

Data set B will be used to begin to address these questions. Two pieces of *a priori* information are available. First, all samples in data set B are "pure" samples of a single ice-type. Since the data set contains no ice-type mixtures, the a_{ik} coefficients should have values close to 0 or 1 if the linear mixing model is appropriate for this texture data. The second useful fact is that the number of end members is expected to be three (first-year, second-year, and multiyear ice). It is not certain that the data analysis will actually support exactly three end members just because conventional arctic wisdom considers these three ice-types as unique. However, K should certainly be small. If, for example, it was found that Eq. (VI-1) could be adequately approximated only if $K > 10$, then the unmixing model is obviously a bad choice for analyzing these KRMS image textures.

The number of end members is associated with the true dimensionality of the data set. In Figure (V-1), which illustrated two end-member mixing, mixtures of the two end members were shown to lie in a straight line connecting the end members. Therefore, in that case, two end members resulted in a one-dimensional mixture data set. The "true" dimensionality of the data set was one less than the number of end members. This relationship holds true for higher

dimensioned feature spaces as well. That is, the number of end members required to approximate Eq. (VI-1) is one larger than the true dimensionality of the data set. Therefore, the search for the required value of K translates into an analysis of the data set dimensionality. Further information on selection of the proper number of end members is available in Bezdek (1974), Miesch (1976b), and Full *et al.* (1981).

B. TRUE DIMENSIONALITY OF DATA SET B

The previous sections showed that the 12 variables of data set B are highly correlated. Therefore, we know that the true dimensionality of the data is less than 12. However, analytic determination of this dimensionality is not a simple problem. The general approach to the structure of multivariate data sets is to assume that the observable, or *manifest*, variables can be represented as functions of a smaller number of *latent*, or hidden, variables. Normally the functional relationships between *manifest* and *latent* variables are constrained to be linear. That is, the *latent* variables are expressed as a weighted sum of the *manifest* variables. If \vec{X}_i is the original data vector whose components are the values of the *manifest* variates, and if \vec{X}_i' is the same data vector expressed in a feature space whose basis vectors are the *latent* variables, then the linear relationship between \vec{X}_i and \vec{X}_i' can be written as a matrix equation.

$$\vec{X}_i' = \vec{X}_i T, \quad (VI - 3)$$

where T is a square matrix containing the coefficients of the linear relationship. Eq. (VI-3) is the standard equation for transformation of coordinate systems. The search for *latent* variables is therefore nothing more than a search for a particular type of coordinate transformation. The number of "significant"

latent variables in the new transformed feature space is effectively the true dimensionality of the original data.

The problem is, then, to find a "best" coordinate transformation matrix, T , to use to in Eq. (VI-3) to convert data vectors from *manifest* to *latent* feature space. One could specify a transformation that requires the variance in the original data set to be preserved in the *latent* feature space. This approach has been given the name Principal Components Analysis (PCA). However, one might also define T so that the original data set can best be reconstructed from a reduced set of the *latent* vector components. This approach has been termed Q-mode Factor Analysis. Other alternatives for defining T could also be proposed. However, we will confine discussion here to the two common techniques just mentioned.

1. Principal Components

Under PCA the original *manifest* variables are transformed (linearly) into a smaller set of new *latent* variables (called principal components), which account for the largest possible amount of variance in the data set. The variance of a multivariate data set is expressed by the covariance matrix, C .

$$C = \frac{1}{(N-1)} \sum_{i=1}^N (\vec{X}_i - \vec{\mu})(\vec{X}_i - \vec{\mu})^T \quad (VI-4)$$

Use of the covariance matrix to define a variance preserving coordinate system transformation relies on the concept of spectral decomposition of a matrix (Johnson and Wichern, 1982). Spectral decomposition of a $K \times K$ symmetric matrix, A , states that the matrix can be expressed as a weighted sum of vector cross-products.

$$A = \sum_{k=1}^K \lambda_k \vec{e}_k \vec{e}_k^T \quad (VI-5)$$

The vectors \vec{e}_k are usually normalized to unit length, i.e., $\vec{e}_k \vec{e}_k^T = 1$, and the λ_k coefficients adjusted accordingly so that the λ s carry the magnitude information for each term of Eq. (VI-5). It is customary, then, to arrange the terms of the equation in order of decreasing values of λ_k . Proof of Eq. (VI-5) can be found in Dunn (1928).

Pairs of λ_k and \vec{e}_k that satisfy Eq. (VI-5) can be found from eigen analysis. Eigenvalues, λ_k , can be found by solving the k^{th} degree polynomial equation (known as the characteristic equation)

$$|A - \lambda I| = 0. \quad (VI - 6)$$

The left side is the determinant of the matrix $A - \lambda I$, where I denotes the identity matrix. Corresponding to every eigenvalue, λ_k , is an eigenvector, \vec{e}_k , such that

$$(A - \lambda_k I) \vec{e}_k^T = 0, \quad k = 1, 2, \dots, K. \quad (VI - 7)$$

Further details concerning eigen analysis are available in numerous references such as Cullen (1966), Finkbeiner (1966), MacDuffee, (1943), Thrall and Tornheim (1957), and White (1966).

Certain facts from eigen analysis will be recalled. One theorem on this subject states that if all the elements of the matrix are positive, which is the case for a covariance matrix, then the eigenvalues are also positive. A second theorem states that for a real symmetric matrix (again, the covariance matrix qualifies) the eigenvectors corresponding to distinct eigenvalues are orthogonal. Yet another useful fact is that, since the eigenvectors in Eq. (VI-5) appear only in quadratic terms, there is a sign ambiguity in eigenvectors. That is, \vec{e}_k and $-\vec{e}_k$ are equivalent eigenvectors. Although Eqs. (VI-6) and (VI-7) provide a means of calculating eigenvalues and eigenvectors, direct solution of these equations is not

feasible when the number of variables is large. For this reason efficient numerical methods have been developed for eigen analysis. The CABFAC computer code used in this study employs Householder's method (Wilkinson, 1960).

When eigen analysis has been performed on the covariance matrix, the columns of the *manifest* to *latent* transformation matrix, T , are formed from the eigenvectors (the principal components) of the covariance matrix. When the data set is then transformed into the *latent* coordinate system (Eq. (VI-3)), the relative magnitudes of the eigenvalues express the proportion of the original variance captured by the new axes in the principal component feature space. Mathematically, the proportion of total variance associated with each principal component is

$$\% \text{ variance in } \vec{e}_k = \frac{\lambda_k}{\text{Tr } C}. \quad (\text{VI} - 8)$$

The application of PCA to data set B results in the proportioning of variance between the 12 principal components shown in Table VI-1. The dimensionality of this data set is not obvious, based on this table. If 90% of the variance is a sufficient representation of the original data set, then the data set is two-dimensional (three end members) as we expect it to be from *a priori* knowledge. However, on what basis can we say that 90% of the variance is sufficient? There is no logical criterion upon which to make this decision. The absence of meaningful criteria for determining what proportion of the variance should be retained is a weakness of PCA.

Another weakness of PCA is the failure to make a distinction between variance and information (Ehrlich and Full, 1987). In addition to ice-type information, variance includes random uncorrelated fluctuations, such as instrument noise, and other fluctuations not related to ice type, such as atmospheric path-

TABLE VI-1. EIGENVALUES OF C FROM DATA SET B

k	Eigenvalue	% Variance	Cum. %
1	33.881	67.76	67.76
2	7.879	15.76	83.52
3	2.869	5.74	89.26
4	2.087	4.17	93.43
5	1.225	2.45	95.88
(6 \rightarrow N)	-	4.12	100.00

length effects. In some data sets the larger principal components may actually capture noise or other aspects of the variance structure not related to ice type. Clearly, it is possible that the third or fourth principal components may carry the most ice-type information, even though they may represent only a small percentage of the total variance. Therefore, determining data set dimensionality from PCA is beset with difficulty.

2. Q-mode Factor Analysis

Miesch (1976b) used a more definitive method for judging dimensionality of a multivariate data set. He proposed that if the dimensionality of a transformed data set is sufficient to adequately represent the original data, that fact would be exhibited by an even distribution of the unaccounted variance across all of the originally measured variables. A concentration of the unaccounted variance in only a few of the original variables can therefore be interpreted as an indication that the reduced data set is of insufficient dimensionality.

The distribution of the residual variance can be determined by back-calculating the sample vectors from eigenvector space to the original variable space. A coefficient of determination between each of the original variable values and the back-calculated values can then be computed. If the coefficients of determination are large and are nearly equal for all of the original variables, then the reduced eigenvector feature space adequately represents the original data according to the Meisch criterion. If one or more of the variables have small coefficients of determination relative to the other variables, then the residual variance is concentrated in these variables and the dimensionality of the reduced feature space is too small to adequately represent the original data.

Since the Meisch criterion for dimensionality involves reconstruction of

the original sample vectors from the reduced sample vectors, an alternative to the covariance matrix as the basis for the eigen analysis is appropriate. If we form a matrix consisting of the cosines between all possible pairs of sample vectors and subject this matrix to eigen analysis, then the result is a dimension-reducing transformation that preserves angular relationships between sample vectors, *i.e.*, preserves maximum reconstructability of the original data set. This type of transformation is consistent with the back-calculating criterion that is used to judge dimensionality. Such a matrix consisting of angular cosines is called a cosine theta, or similarity matrix. The use of the cosine theta matrix for eigen analysis, rather than PCA based on the covariance matrix, would seem to be the better choice because of the resulting maximization of back-calculation accuracy. Eigen analysis of the cosine theta matrix is commonly called Q-mode Factor Analysis.

The cosine theta matrix is an $N \times N$ matrix, where N is the number of samples in the data set. Such a matrix would require large storage space and long eigen analysis computation times for large data sets. However, it is not necessary to actually calculate and analyze the cosine theta matrix. Klován and Imbrie (1971) show that only m of the eigen values of the cosine theta matrix have non-zero values, where m is the dimensionality of the original feature space. Furthermore, the non-zero eigenvalues and their associated eigenvectors have been shown to be identical to those of the $m \times m$ cross-product matrix, P .

$$P = \sum_{i=1}^N \bar{X}_i^T \bar{X}_i \quad (VI - 9)$$

Thus in this application, P is the mathematical equivalent of the cosine theta matrix and P , whose storage requirements are independent of data set size, is utilized in place of the cosine theta matrix for Q-mode Factor Analysis in the

Extended CABFAC computer code (Klovan and Imbrie, 1971; Full *et al.*, 1981). Note the similarity between the covariance matrix (Eq. (VI-4)) and the cross-product matrix (Eq. (VI-9)). The only difference is the subtraction of the mean vector in the case of the covariance matrix. For the special case of data sets with zero means, Principal Components Analysis and Q-mode Factor Analysis are identical.

Eigen analysis of the cross-product (cosine theta) matrix for data set B results in the eigenvalues shown in Table VI-2. The coefficients of determination between original and back-calculated sample vectors are plotted as a function of number of end members in Figure VI-1. The figure shows that for two end members, two of the variables (LM4 and LM5) have coefficients of determination of 0.4073 and 0.0476 compared to coefficients of 0.8 or larger for the other 10 variables. According to the Miesch criterion, this distribution of residual variance indicates that two end members are not adequate. As we move to three end members, the coefficients of determination rise to 0.8 or higher for all variables, thus indicating that three end members are sufficient to linearly model data set B. This result is consistent with the *a priori* knowledge of the ice-texture data set structure. We will therefore proceed into unmixing calculations, considering data set B to be two-dimensional, i.e., the data will be interpreted as linear mixtures of three end members. The reduced feature space must therefore be three-dimensional in order to accommodate a two-dimensional data set that is constrained to constant sum.

The matrix, T , which transforms the original twelve-dimensional data into a three-dimensional feature space using Eq. (VI-3), is constructed from the first

TABLE VI-2. EIGENVALUES OF P FROM DATA SET B

k	Eigenvalue	$\frac{\sum \lambda_k}{\text{Tr } P}$
1	40.76	81.52
2	7.10	95.72
3	0.96	97.64
4	0.43	98.49
5	0.30	99.08
(6 \rightarrow N)	-	100.00

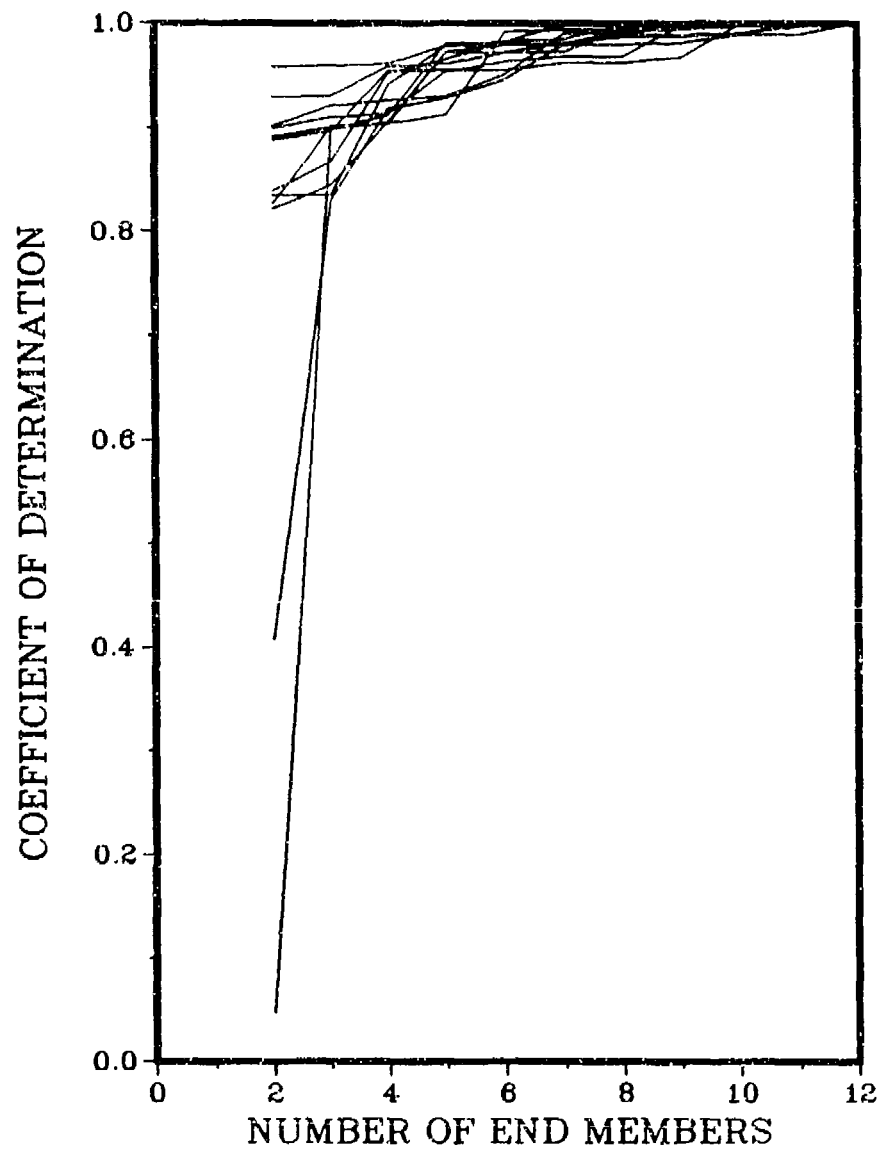


FIGURE VI-1. Coefficients of determination for the 12 variables of data set B as a function of number of end members.

three eigenvectors of P .

$$T = \begin{bmatrix} 0.170 & 0.355 & -0.079 \\ 0.398 & -0.229 & 0.028 \\ 0.171 & 0.357 & -0.012 \\ 0.321 & -0.235 & 0.023 \\ 0.303 & 0.407 & -0.096 \\ 0.391 & -0.249 & 0.227 \\ 0.184 & 0.368 & 0.120 \\ 0.396 & -0.197 & -0.138 \\ 0.143 & 0.223 & 0.505 \\ 0.292 & 0.002 & -0.737 \\ 0.201 & 0.360 & 0.086 \\ 0.314 & -0.240 & 0.305 \end{bmatrix} \quad (VI - 10)$$

C. THE VARIMAX TRANSFORMATION

The CABFAC software performs an additional rotation on the data vectors in the reduced feature space (Klovan and Imbrie, 1971). The procedure used is the well-known varimax method of Kaiser (1958). The varimax rotation is a rigid, orthogonal rotation that does not change the relationships between sample vectors. The varimax rotation is applied simply as a matter of convenience in interpretation. The property of the varimax rotation that is of interest here is the fact that the rotation tends to place the data cloud in the positive orthant. This result is highly desirable if one wishes to plot the data for visual analysis and interpretation as we have done in the following sections.

The result of applying the varimax criterion to the reduced feature space data is a rotation matrix, V , which for the reduced three-dimensional version of data set B, was found to be

$$V = \begin{bmatrix} 0.812 & 0.573 & 0.115 \\ -0.579 & 0.815 & 0.028 \\ 0.078 & 0.089 & -0.993 \end{bmatrix}. \quad (VI - 11)$$

To transform data directly from the original feature space into the rotated, reduced feature space, one can combine transformation matrices resulting from

Factor Analysis and the varimax rotation into a single transformation matrix by taking the matrix product of T and V .

$$TV = \begin{bmatrix} -0.087 & 0.378 & 0.104 \\ 0.456 & 0.061 & 0.010 \\ -0.083 & 0.385 & 0.038 \\ 0.398 & 0.009 & 0.006 \\ -0.015 & 0.498 & 0.137 \\ 0.477 & 0.057 & -0.190 \\ -0.070 & 0.413 & -0.091 \\ 0.423 & 0.071 & 0.174 \\ 0.013 & 0.305 & -0.482 \\ 0.177 & 0.117 & 0.763 \\ -0.054 & 0.414 & -0.057 \\ 0.416 & 0.024 & -0.276 \end{bmatrix} \quad (VI-12)$$

Equation (VI-12) is the matrix that will be applied to transform data set B into a three-dimensional feature space for the unmixing analysis that follows.

D. END-MEMBER PROPORTIONS IN DATA SET B

Knowing from factor analysis that data set B can be modeled using three end members, and having established a matrix to transform the image texture data into a three-dimensional feature space, we now proceed to the unmixing of the individual samples. In other words, we calculate the a_{ik} coefficients of Eq. (VI-1).

First it will be instructive to plot data set B as it exists in the reduced feature space. Several issues crucial to unmixing strategy will become apparent from a visual inspection of the data. The reduced feature space is three-dimensional, but the constant-sum constraint causes the data points to fall in a plane rather than throughout the three-dimensional volume. The planar nature of the data allows for easy plotting in two dimensions, as shown in Figure VI-2, where F_1 , F_2 , and F_3 are the basis vectors of the reduced feature space. The feature space has

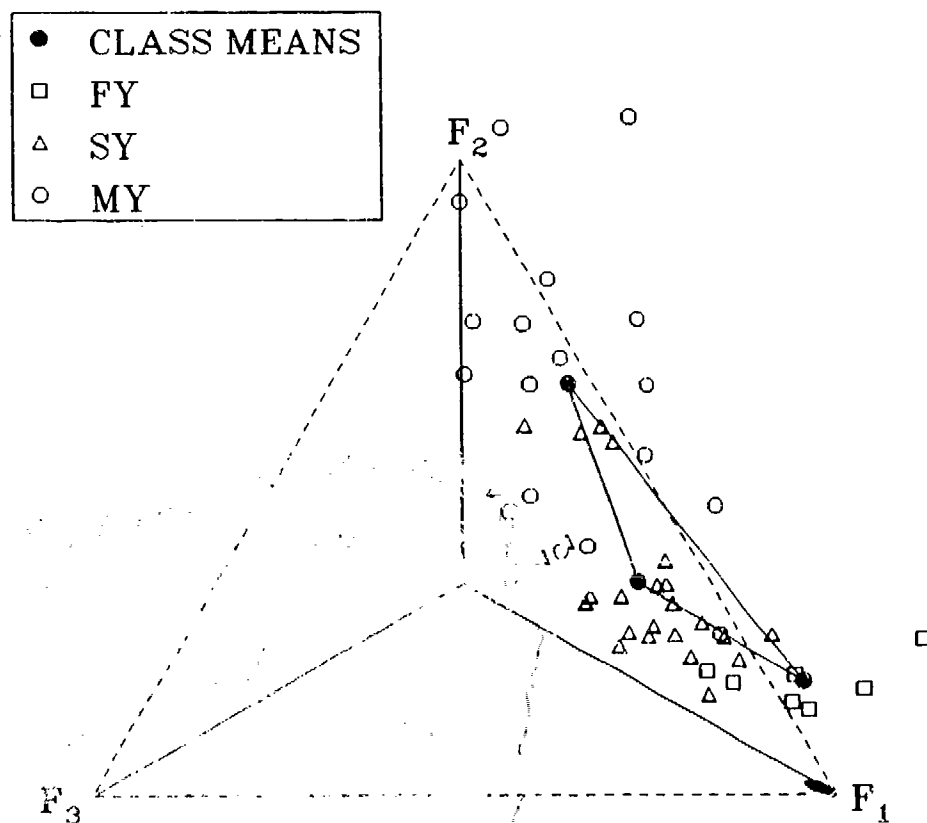


FIGURE VI-2. Plot of data set B in the reduced feature space. Solid dots indicate mean positions for samples belonging to FY, SY, and MY ice classes.

been rotated for plotting so that the $F_1 = F_2 = F_3$ line is perpendicular to the page. The constant-sum plane in Figure VI-2 therefore lies in the plane of the page.

The plot denotes first-year, second-year, and multiyear samples with different symbols. Also plotted are the mean vectors for each of the ice classes. These means will be assumed to be the appropriate end members for these initial unmixing calculations. Since we know ice type for each of these samples, the end-member index, k , can be associated with an ice-type. Therefore, the end-member notation \vec{E}_{FY} , \vec{E}_{SY} , and \vec{E}_{MY} will be used. The numerical values for end members are (i.e., class means) as follows.

$$\begin{aligned}\vec{E}_{FY} &= [0.868 \quad 0.177 \quad -0.045] \\ \vec{E}_{SY} &= [0.567 \quad 0.336 \quad 0.097] \\ \vec{E}_{MY} &= [0.320 \quad 0.648 \quad 0.032]\end{aligned}\tag{VI-13}$$

Pairs of end members are connected in Figure VI-2 to indicate areas of linear unmixing between these pairs. These end-member connecting lines form a geometric figure (a triangle in this case) that is frequently referred to as the mixing polytope. Samples falling on an edge of the polytope represent mixtures of the two end members that form that line segment, with no contribution from any other end members. Sample points falling within the area of the polytope represent mixtures of three or more end members. According to linear mixing theory, samples that fall outside the polytope should not occur because they represent mixing proportions greater than one or less than zero, which is physically impossible. However with real data, samples do exist outside the mixing polytope. Samples outside the polytope can result from random noise in the data, or from improper selection of the number or positions of end members.

Two important observations are made from Figure VI-2. First, the data are noisy. The scatter in the sample points belonging to a single ice-type is larger than the size of the mixing polytope. Second, the polytope is a highly elongated triangle. The SY end member falls very close to the midpoint of the FY/MY mixing line. That is, it will be very difficult to differentiate between SY ice and an equal mixture of FY and MY. Each of these two factors will contribute to rather large errors when sample proportions calculated from by unmixing this data are compared with actual image compositions. Any steps that could be taken to reduce noise levels in the data, or to expand the polytope to more closely approximate an equilateral triangle, would improve the unmixing results. However, before investigating these possibilities, let us calculate the unmixing accuracies associated with the data shown in Figure VI-2.

1. Unmixing Data Set B

The computer code FUZZY QMODEL (Full *et al.*, 1982) is used to calculate mixing proportions. The main function of the QMODEL series of programs (Klovan and Miesch, 1976; Miesch, 1976, and Full *et al.*, 1981) is to determine end members by one of several optional methods. Since the present interest is in using class means as end members, these aspects of QMODEL are not needed at this time. The background and detailed description of QMODEL are therefore deferred until a later section where QMODEL's end-member determination capabilities are employed. For now it suffices to say that we employed a FUZZY QMODEL option that allows vertices of the mixing polytope to be supplied as input to the program. This option utilizes only that portion of QMODEL which calculates end-member proportions for each sample based on its position with respect to the supplied polytope.

The compositions of the samples in data set B are known. Each sample has a 1.0 proportion of one ice type and a 0.0 proportion of the other two ice types. Each sample, therefore, provides three numbers (proportions of FY, SY, and MY) to indicate unmixing accuracy. Thus the 50 samples in the data set provide 150 measures of accuracy. Looking at the 50 cases, which should have resulted in proportions of 1.0, QMODEL-calculated proportions were found to average 0.985 and had a standard deviation of ± 0.643 about that value. QMODEL calculated a mean proportion value of 0.118 ± 0.615 for the 100 cases where 0.0 was the correct answer. The calculated proportions where 0.0 is the expected answer are significantly biased toward the positive side. Combining all 150 error values gives a single rms proportion error of 0.632 for data set B

These unmixing results are very poor when one considers that by randomly assigning each sample to a class, an rms proportion accuracy of 0.67 can be achieved. Unmixing of the data in Figure VI-2 based on the polytope shown is, therefore, only slightly superior to random guessing. However, the situation is not as grave as it first seems. The global approach to image texture analysis was undertaken because we concluded intuitively that ice type could not be accurately determined based on the texture of small neighborhoods. These poor unmixing results supply experimental confirmation of that initial conclusion. This unmixing analysis has demonstrated that even a 64×64 pixel image sample size is not large enough to generate stable ice-texture statistics. As sample size increases, the unmixing accuracy can be expected to improve.

If we assume that the rms error in calculated proportions decreases as $(\text{no. of pixels in sample})^{-1/2}$, then for a full 512×312 image (100 pixels on each edge of a 512×512 image not included for reasons discussed in Section II) the rms error in calculated proportions would be reduced to 0.101. (The assumed

dependence of accuracy on sample size is reasonable, since we have shown that the original variables are normally distributed and since all subsequent operations leading to composition estimates are linear.) This full image error of 10% in ice-type compositions is approaching the useful range for many applications. However, before we are content to lay all of our hopes on increased image size alone, it will be worthwhile to investigate some possibilities for reducing the noise in the data and for improving the shape of the mixing polytope. This problem has consumed a great deal of effort, the details of which add no value to this dissertation. However, some results will be presented here. The reader should note that these are summary results and that the problem of reducing noise and improving polytope shape are not as straightforward as they might appear from the following brief treatment of these topics.

2. Noise Reduction

A significant portion of the random noise in the data seemed to originate with the LM4 and LM5 variables. This judgment is based on observations summarized as follows. In Figure III-12, which shows the Laws' Mask images for a typical KRMS image of multiyear ice, it does not appear visually that masks 4 and 5 contain much ice-type information. In Table IV-1 both LM4 and LM5 were found to be so devoid of second-year /multiyear discrimination power that no solution existed for a Bayesian threshold to separate these ice types. These variables seemed to be useless until it was noted in Section IV that when pairs of texture measures were considered in a bivariate classifier, the MBRL-LM5 combination gave better ice-type classification than any other pair of variables. It therefore seemed wise to include LM4 and LM5 in data set B. Is it a mistake to include these variables? Can the noise in the data set be reduced and ice-type

proportion accuracy be improved if these variables are excluded?

The variables LM4 and LM5 were removed from data set B and the data were reprocessed through the SIFT, the CABFAC, and the QMODEL programs. Figure VI-3 is a plot of the resulting data set in the reduced feature space. The noise has been reduced by removal of the two variables in question. The reduced noise level is especially apparent for first-year ice. Note in Figure VI-3 that the square symbols are closely grouped about their mean point, whereas in Figure VI-2, the square symbols are more widely scattered. The LM4 and LM5 variables are apparently responsible for the scatter of first-year ice points in Figure VI-2 and for increased noise in the other ice types as well. Even though elimination of these two variables has reduced the noise, the overall result is not desirable. The third dimension of the polytope has also been lost. The second-year end member now lies almost exactly on the first-year/multiyear edge of the polytope. We have therefore eliminated our ability to distinguish second-year ice in order to achieve a reduction in noise levels.

This result is not surprising, since the plot of coefficients of determination in Figure VI-1 showed us that LM4 and LM5 were required for a three-end-member fit to the linear mixing model. We therefore must retain LM4 and LM5, in spite of their noisy nature, if the ability to unmix KRMS images into three ice types is to be retained. The rms error in calculated ice-type proportions for the data shown in Figure VI-3 is 1.916, which is larger than the rms error of 0.632 by a factor of three when LM4 and LM5 are included.

3. Polytope Shape Improvement

The other possibility previously mentioned for improving unmixing accuracy is to alter the shape of the mixing polytope from that shown in Figure

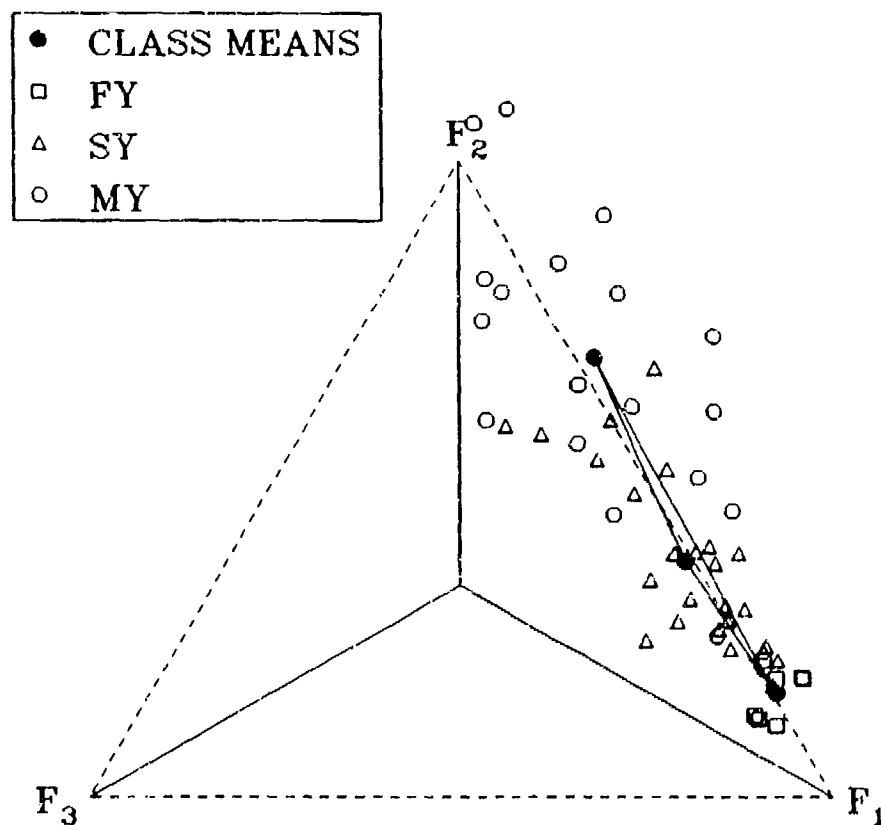


FIGURE VI-3. Plot of data set B in the reduced feature space with variables LM4 and LM5 excluded from the analysis.

VI-2. In other words, can some change be made such that the second-year end member is not as close to the midpoint of the FY/MY side of the polytope as is the case in Figure VI-2? One way to do this change is to include the MEAN variable, along with the 12 texture variables in data set B. To support this idea, the author recalls attention to Table IV-1, where MEAN was shown to be a significantly superior ice-type classifier to any of the individual texture measures. Pairwise combinations of MEAN and a texture variable resulted in ice-type classification accuracies as high as 98% compared to 86% for the best combination of two texture variables in Table IV-2. Therefore, one would expect the addition of MEAN to the texture data set might improve classification accuracy by improving the shape of the polytope. MEAN had previously been excluded from consideration because it is not a true texture measure. However, in view of the relatively poor mixing performance demonstrated by data set B containing texture information alone, it is appropriate to investigate what performance can be obtained by mixing MEAN with texture measures.

The ability to incorporate nontexture variables into the analysis is one advantage of the approach employed here. The treatment of multivariate samples as vectors in a feature space, combined with global statistics, and the unmixing of the global values, is a paradigm that allows complete flexibility in combining different variable types into a unified analysis.

Figure VI-4 shows the results of a 13-variable analysis (12 texture measures plus mean brightness). The mixing polytope in Figure VI-4 is much better shaped than that in Figure VI-2. Unmixing statistics based on this polytope show an rms proportion error of 0.341, which is significantly better than the value of 0.632 obtained from texture measures alone. When extrapolated to a 512×312 pixel sample size, the rms error for combined texture and mean brightness is

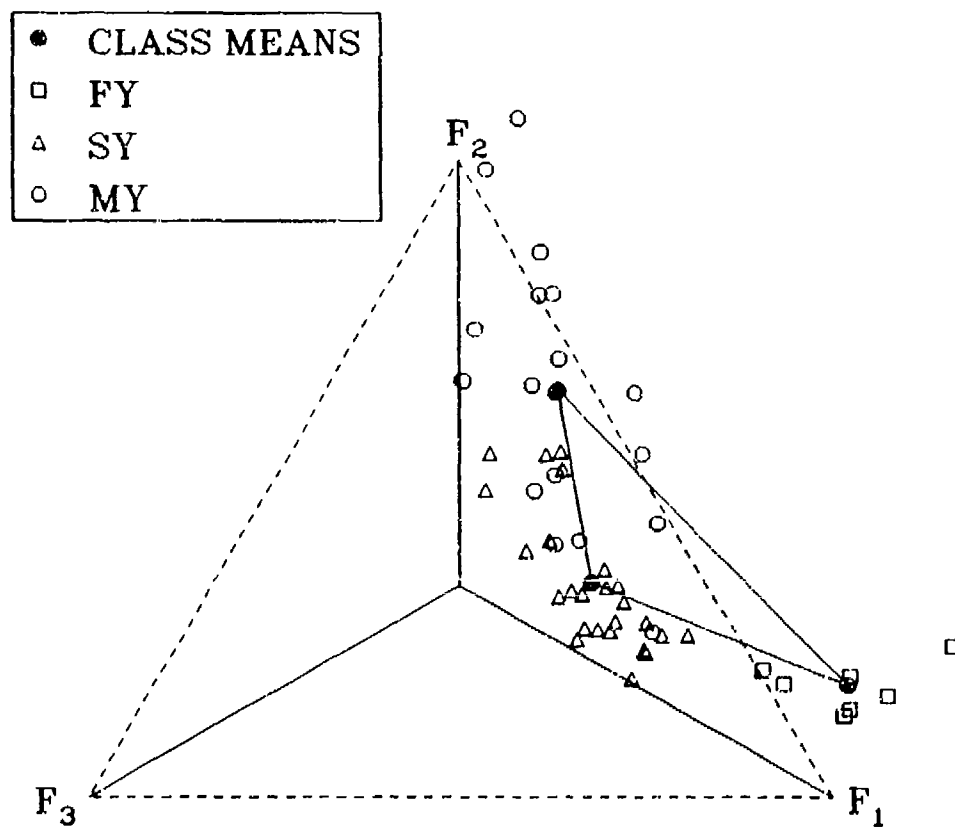


FIGURE VI-4. Plot of data set B in the reduced feature space with MEAN included along with the 12 original texture variables.

0.0055.

4. Combined Noise Reduction and Polytope Improvement

Figure VI-3 showed that removing LM4 and LM5 from the analysis removed much of the noise in the data set, but that fact could not be taken advantage of because these two variables were absolutely necessary to form a second-year end member. Figure VI-4 then showed that using MEAN along with texture measures enhances the separability of the second-year end member even better than did LM4 and LM5. Therefore, the question arises: can one utilize MEAN to provide second-year separability and at the same time remove LM4 and LM5 to reduce noise? Figure VI-5 shows this case. The data plotted in this figure are derived from 10 texture variables (data set B minus LM4 and LM5) plus MEAN. The rms error in proportion calculations is 0.334. This error is only slightly better than the error in Figure VI-4, where LM4 and LM5 are included. Nevertheless, the error is smaller and the combination of removing LM4 and LM5 while adding MEAN is the best ice-type proportion estimator of those evaluated.

We will therefore proceed from this point with two data sets to be evaluated in parallel. The pure texture case, which includes the 12 texture measures of data set B, will be pursued to illustrate what can be done with texture information only. Then a second case consisting of 10 texture measures (LM4, LM5 removed) and MEAN will be pursued in parallel to illustrate improved ice-type classification accuracy through combination of texture with other information.

If the combination of texture and mean intensity is to be carried forward into further analysis, then it is necessary to confirm that a three-end-member approximation to the mixing model (Eq. (VI-1)) is still appropriate after having dropped two texture variables and added the mean intensity. To confirm this fact,

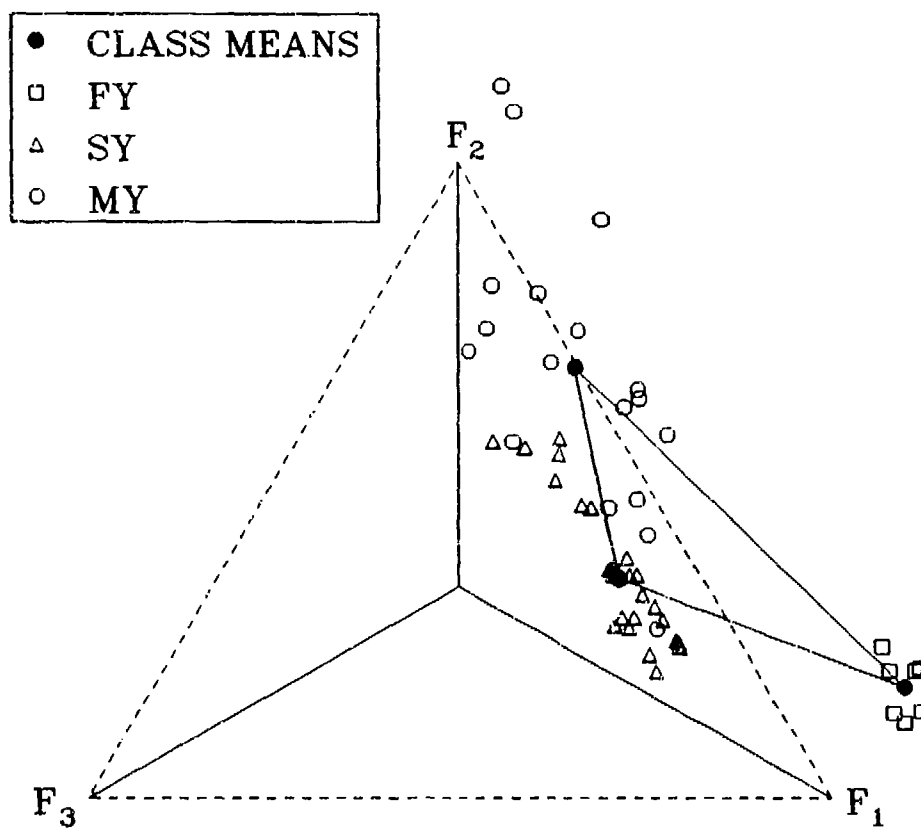


FIGURE VI-5. Plot of data set C in the reduced feature space.

the coefficients of determination calculated in CABFAC are plotted in Figure VI-6 for the texture/mean combination case. Here, as in the texture-only case, it can be seen that an even distribution of residual variance results for three end members or more. Therefore, the three-end-member solution is still justified for this combined case.

For sake of completeness the data set, which consists of 10 texture variables and MEAN, is tabulated in Appendix E and is given the name data set C. The transformation matrix to go from the data set C eleven-dimensional feature space to the reduced three-dimensional varimax space is as follows.

$$TV = \begin{bmatrix} -0.091 & 0.389 & -0.085 \\ 0.453 & 0.056 & 0.086 \\ -0.090 & 0.394 & -0.130 \\ 0.397 & 0.001 & 0.110 \\ -0.028 & 0.504 & 0.034 \\ 0.473 & 0.058 & -0.217 \\ -0.086 & 0.406 & -0.086 \\ 0.439 & 0.104 & -0.213 \\ -0.064 & 0.423 & -0.260 \\ 0.408 & 0.009 & -0.169 \\ 0.161 & 0.279 & 0.872 \end{bmatrix} \quad (VI - 14)$$

The numerical values for end members (i.e., class means) are as follows.

$$\begin{aligned} \vec{E}_{FY} &= [1.010 \quad 0.175 \quad -0.186] \\ \vec{E}_{SY} &= [0.532 \quad 0.358 \quad 0.110] \\ \vec{E}_{MY} &= [0.320 \quad 0.677 \quad 0.003] \end{aligned} \quad (VI - 15)$$

5. Conclusions from Unmixing Pure Samples

Analysis of this set of fifty 64×64 pixel sample images of pure ice types has led to the results summarized below.

- Q-mode Factor Analysis has suggested a three-end-member mixing solution for this data. Three end members is consistent with the number of conventional ice categories (first-year, second-year, and multiyear) associated

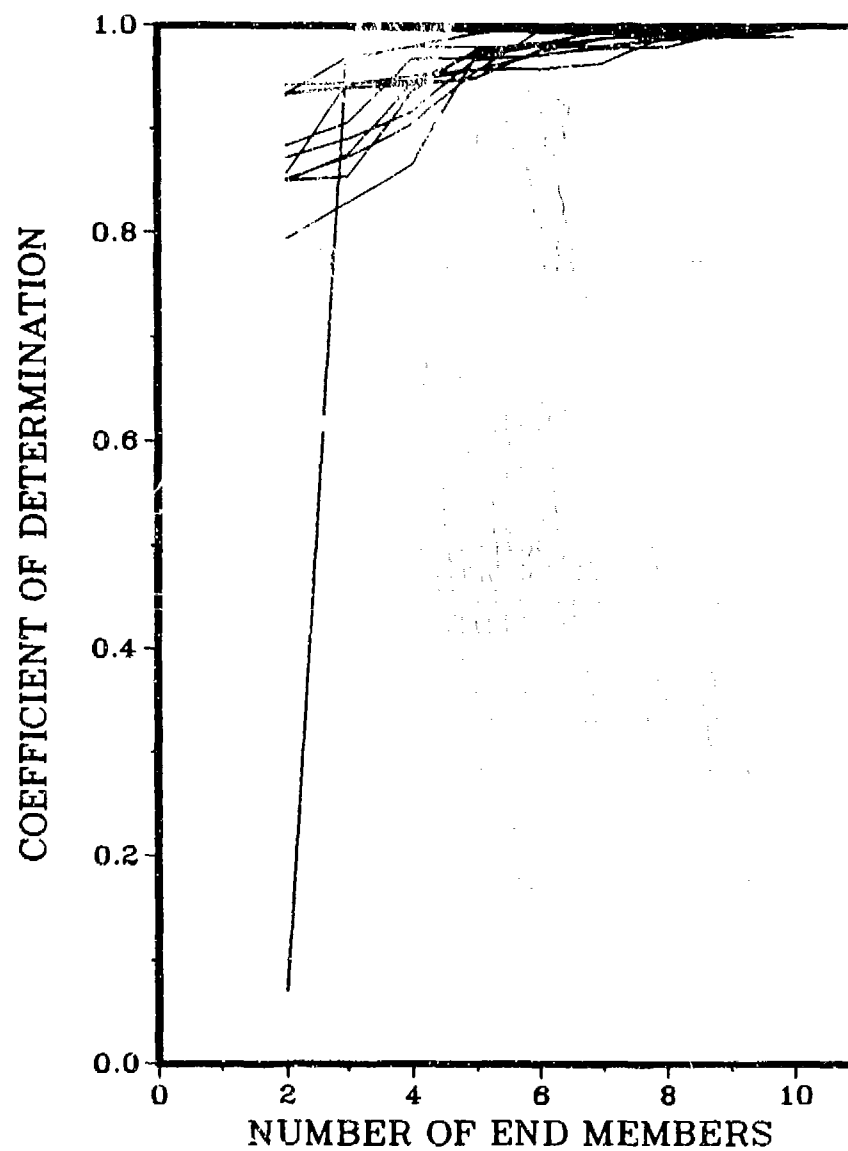


FIGURE VI-6. Coefficients of determination for the 11 variables of data set C as a function of number of end members.

with this data. This offers encouragement that the results of the unmixing mathematics will be easily interpretable in terms common to Arctic science.

- A determination of image composition by ice type based on unmixing has been demonstrated to be accurate to an rms value of 0.632 using only texture variables on images of this size. Accuracy has been estimated to be 0.10 for 512×312 pixel images.
- Accuracy in determination of image composition has been shown to improve by a factor of approximately two (0.334) if mean image brightness level is included along with image texture information (0.055 for 512×312 images).

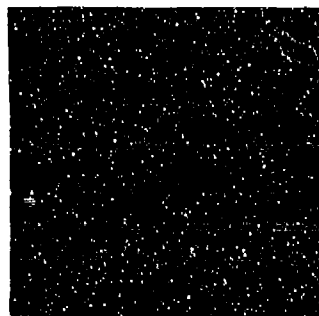
VII. UNMIXING OF SYNTHETIC MIXTURES

A. SYNTHETIC MIXTURE DATA SET

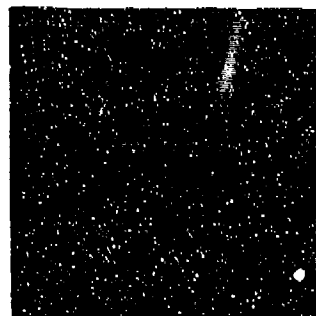
Section IV-D established that many of the 12 image texture measures used in data sets B and C in preceding sections do not mix linearly, unless the mean brightness value is the same in each of the ice-type subareas of the mixture. The first step in the analysis of mixture images is, therefore, to create a mixture data set that meets this requirement of identical mean intensity in each component of the mixture. Synthetic mixtures created in this manner are distinctly different from naturally occurring mixtures where mean intensity varies with ice type. The more general case of varying means will be treated in Section VIII. Here, we address the simpler special case represented by these synthetic mixtures. By working first with synthetic mixtures and then with natural mixtures, it will be possible to evaluate the magnitude of the increase in proportion measurement errors resulting from varying means among the components of the mixture.

Forty-two synthetic mixture images, named data set D, have been created by combining the pure samples of data set A (Figures VI-1 through VI-10) in various ways. The intensity values in the subimages were offset prior to merging so that each has a mean value of 128. The mean-adjusted components are then mosaicked to form 128×128 pixel mixture images. The mixture images, labeled MX1 through MX42, are shown in Figures VII-1 through VII-7. The same 12 texture measures associated with data set B are calculated for these synthetic mixture images. Mean intensity is also calculated, but this calculation is performed prior to the mean adjustment step so that MEAN has the value that would result from mosaicking images without an adjustment. The texture and mean values for the images of data set D are tabulated in Appendix F.

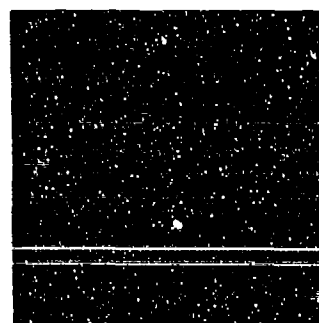
MX1



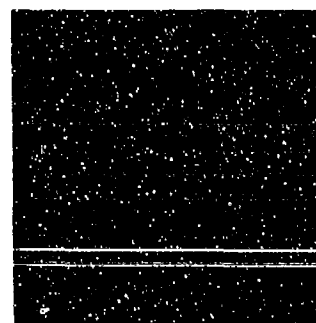
MX2



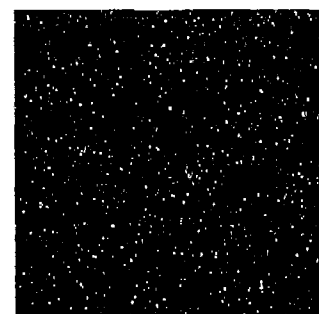
MX3



MX4



MX5



MX6

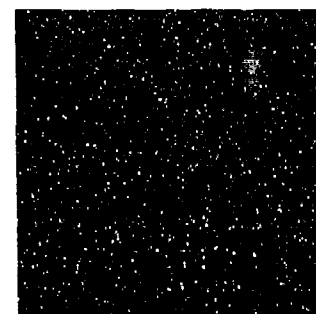
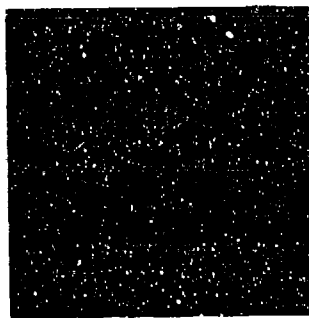
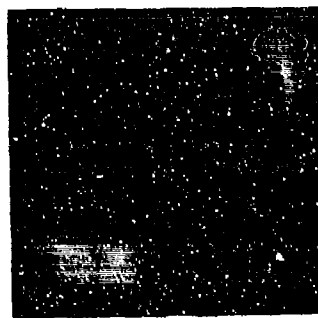


FIGURE VII-1. Synthetic mixture images MX1 - MX6 produced by merging pure image samples from data set A.

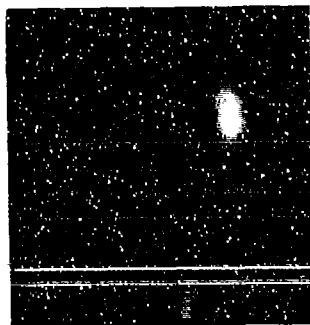
MX7



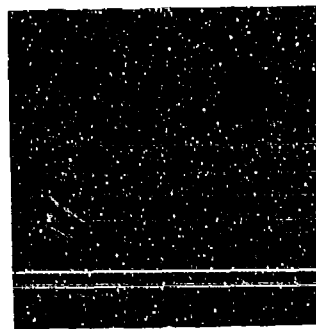
MX8



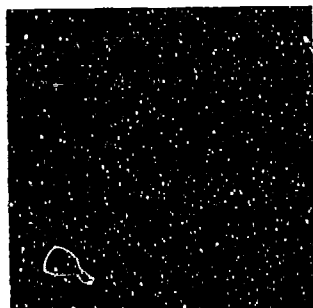
MX9



MX10



MX11



MX12

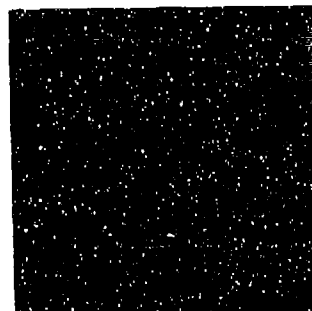
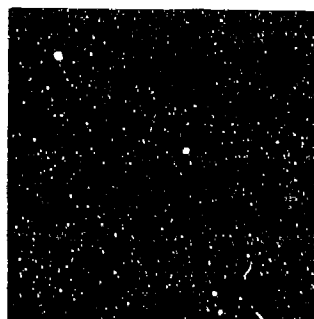
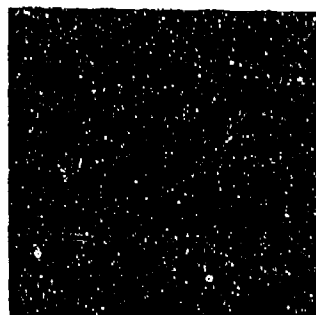


FIGURE VII-2. Synthetic mixture images MX7 -- MX12 produced by merging pure image samples from data set A.

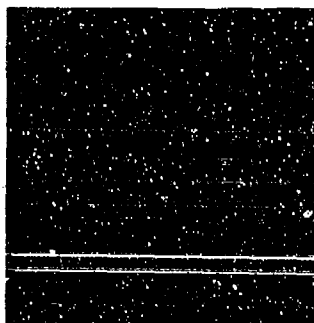
MX13



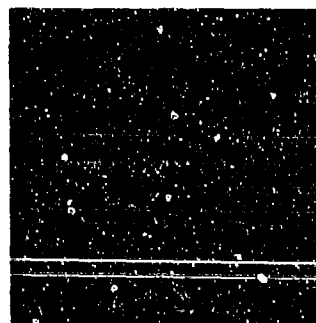
MX14



MX15



MX16



MX17



MX18

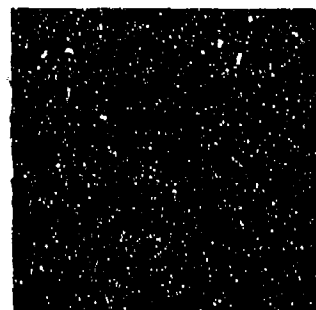
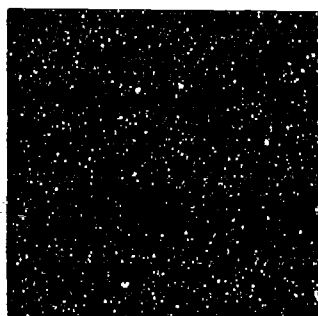
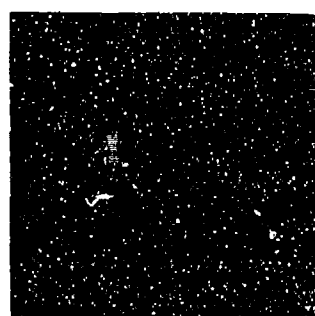


FIGURE VII-3. Synthetic mixture images MX13 -- MX18 produced by merging pure image samples from data set A.

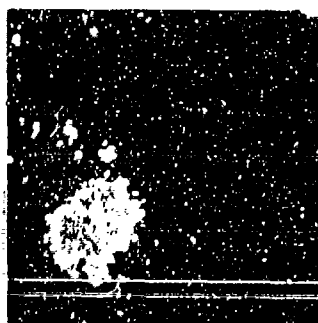
MX19



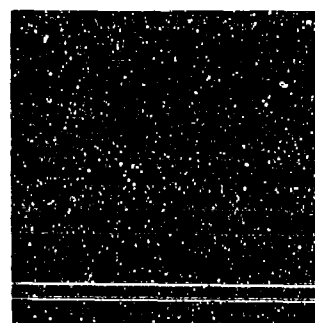
MX20



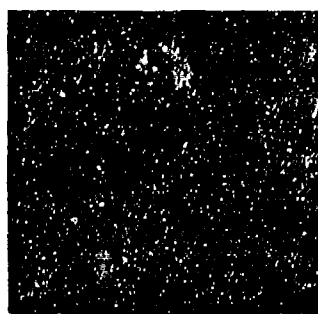
MX21



MX22



MX23



MX24

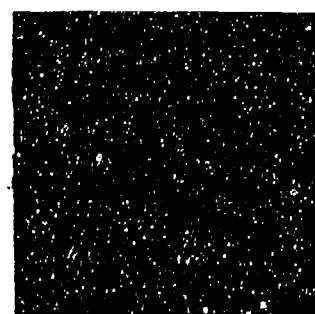
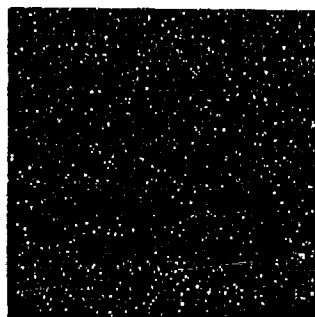
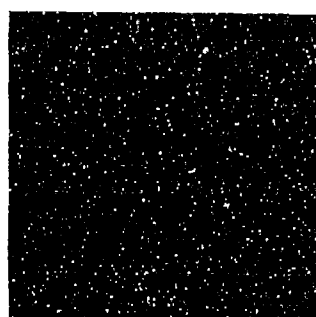


FIGURE VII-4. Synthetic mixture images MX19 - MX24 produced by merging pure image samples from data set A.

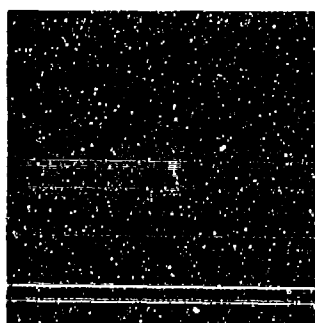
MX25



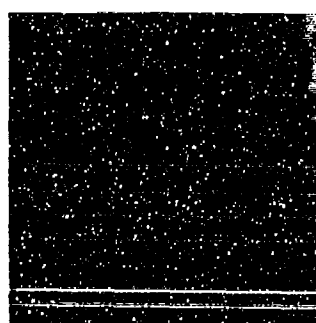
MX26



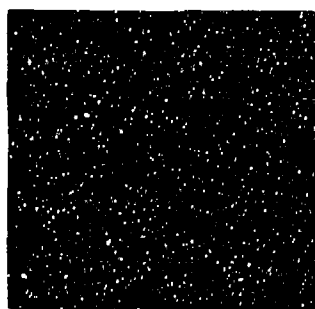
MX27



MX28



MX29



MX30

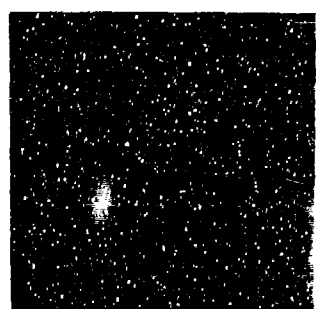
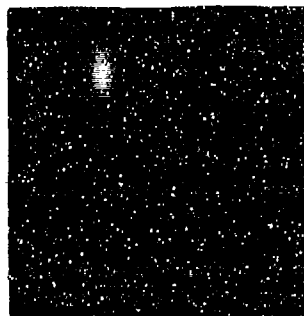
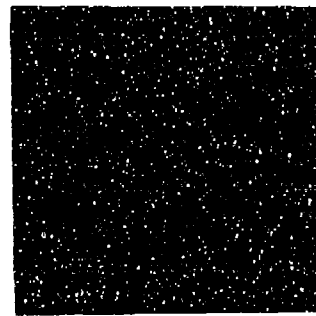


FIGURE VII-5. Synthetic mixture images MX25 - MX30 produced by merging pure image samples from data set A.

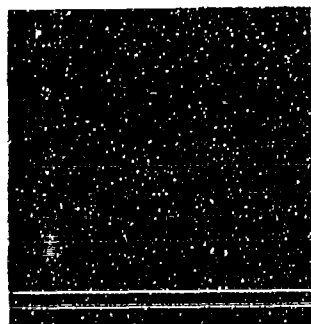
MX31



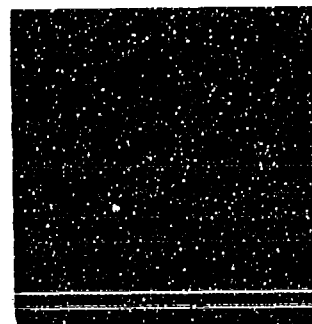
MX32



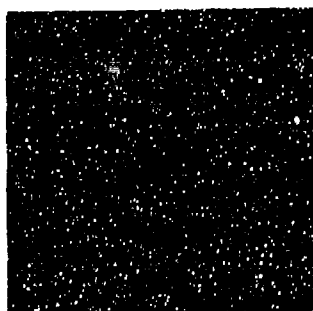
MX33



MX34



MX35



MX36

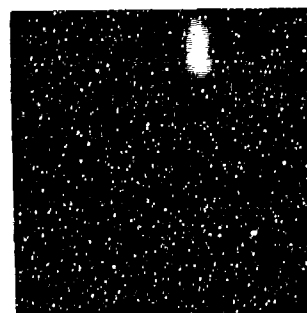
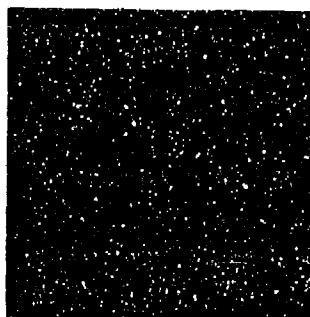
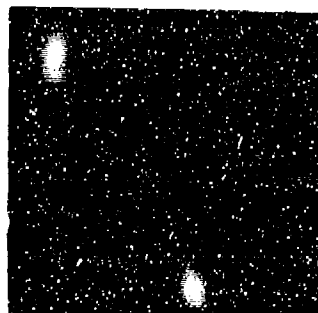


FIGURE VII-6. Synthetic mixture images MX31 - MX36 produced by merging pure image samples from data set A.

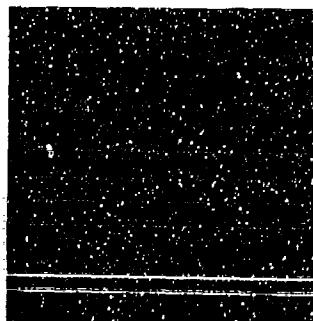
MX37



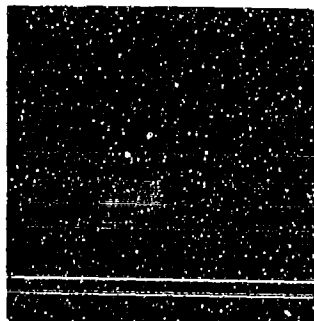
MX38



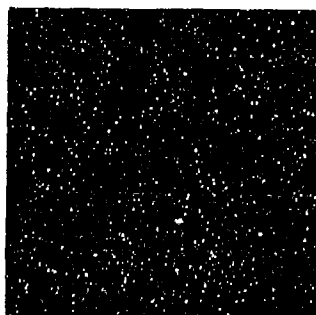
MX39



MX40



MX41



MX42

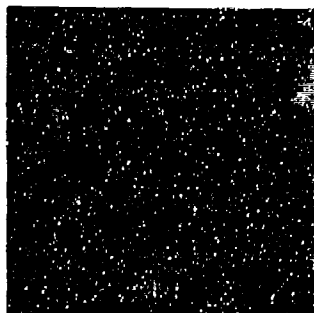


FIGURE VII-7. Synthetic mixture images MX37 -- MX42 produced by merging pure image samples from data set A.

B. UNMIXING WITH CLASS MEANS AS END MEMBERS

1. Texture Measures Only

The 12 selected texture measures for the synthetic mixture images were processed through the SIFT procedure to form data set E given in Appendix G. The CABFAC program was used to find the transformation matrix, TV , which would put data set E into a three-dimensional reduced feature space. The transformation matrix was found to be

$$TV = \begin{bmatrix} -0.060 & 0.412 & 0.093 \\ 0.443 & 0.023 & 0.096 \\ -0.047 & 0.434 & 0.044 \\ 0.428 & 0.012 & -0.005 \\ -0.019 & 0.417 & 0.229 \\ 0.455 & 0.018 & -0.071 \\ -0.034 & 0.386 & -0.005 \\ 0.408 & 0.009 & 0.178 \\ 0.100 & 0.412 & -0.417 \\ 0.085 & -0.003 & 0.820 \\ -0.046 & 0.384 & 0.072 \\ 0.469 & 0.034 & -0.200 \end{bmatrix} \quad (VII - 1)$$

The mean vectors calculated from data set A for each ice class were also SIFTed and transformed into the reduced feature space. Coordinates of the mean vectors, which will serve as end members for this unmixing step, are given below.

$$\begin{aligned} \vec{E}_{FY} &= [0.861 \quad 0.089 \quad 0.050] \\ \vec{E}_{SY} &= [0.463 \quad 0.278 \quad 0.260] \\ \vec{E}_{MY} &= [0.170 \quad 0.674 \quad 0.157] \end{aligned} \quad (VII - 2)$$

Figure VII-8 shows data set E and the class means which have been connected to form the mixing polytope.

The ice-type compositions based on this polytope, which was input to QMODEL for use as end members, are given in Table VII-1. The accuracy (i.e.,

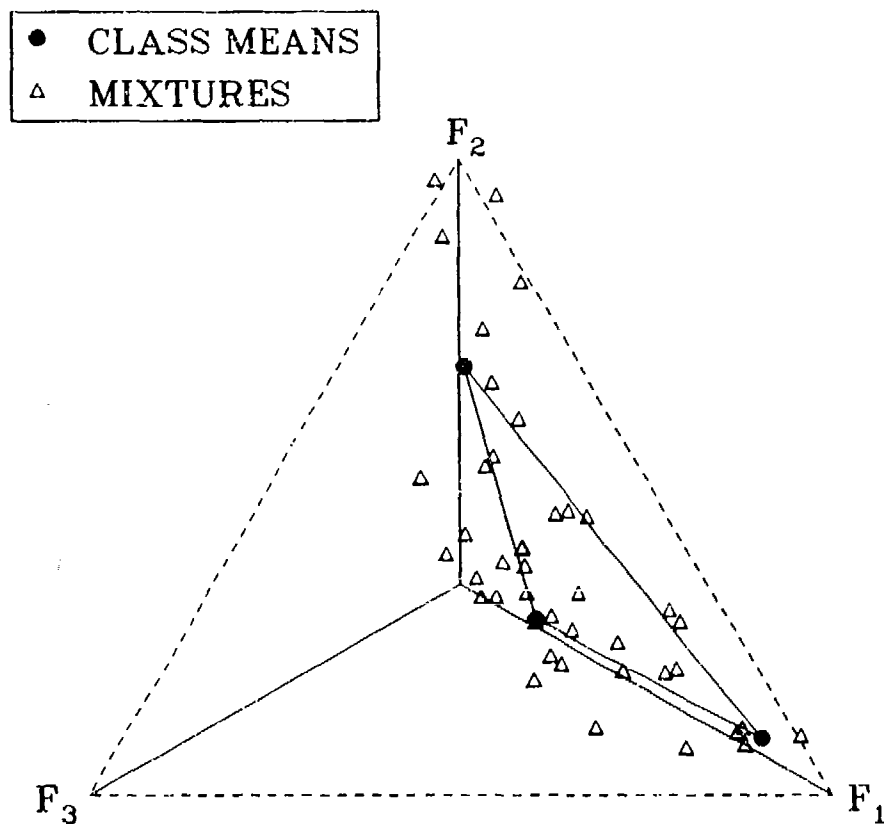


FIGURE VII-8. Plot of data set E in the reduced feature space. Solid dots indicate mean positions for samples belonging to FY, SY, and MY ice classes in data set A.

TABLE VII-1. UNMIXING RESULTS FOR DATA SETS E AND F
Class Means from Data Sets B and C as End Members

Sample	Known Proportions (%)			Set E - Texture Only			Set F - Texture + Mean		
	FY	SY	MY	FY	SY	MY	FY	SY	MY
MX1	0	0	100	35.46	-98.46	163.01	-1.37	-28.58	129.95
MX2	0	0	100	10.65	-72.15	161.50	-14.91	-23.49	138.40
MX3	0	0	100	6.95	-51.09	144.14	-8.73	-19.24	127.96
MX4	0	100	0	-27.97	122.24	5.73	-16.87	104.20	12.67
MX5	0	100	0	18.55	76.84	4.61	5.88	99.75	-5.63
MX6	0	100	0	41.61	47.95	10.44	6.03	112.40	-18.43
MX7	100	0	0	89.59	13.82	-3.41	97.09	0.80	2.11
MX8	100	0	0	92.22	7.69	0.09	99.11	-5.62	6.51
MX9	100	0	0	116.82	-27.76	10.93	100.24	6.91	-7.15
MX10	50	25	25	-0.87	102.27	-1.39	45.33	11.42	43.26
MX11	50	25	25	27.49	47.48	25.03	43.41	16.24	40.34
MX12	25	50	25	40.51	62.27	-2.77	28.93	83.43	-12.36
MX13	25	50	25	4.09	62.18	33.73	9.95	52.24	37.81
MX14	25	25	50	2.38	72.76	24.87	18.23	43.17	38.59
MX15	25	25	50	2.20	29.49	68.31	6.49	21.31	72.20
MX16	50	0	50	8.72	85.66	5.62	56.65	-9.29	52.63
MX17	0	50	50	37.15	-75.91	138.76	-18.91	36.96	81.95
MX18	50	50	0	73.68	-8.77	35.09	46.80	42.34	10.85
MX19	25	75	0	2.11	113.53	-15.64	21.32	75.67	3.01
MX20	75	25	0	92.12	15.18	-7.31	75.57	43.17	-16.73
MX21	75	0	25	69.94	-7.95	38.01	64.67	2.77	32.56
MX22	25	0	75	11.78	-10.15	98.36	19.38	-28.46	109.08
MX23	0	75	25	20.10	-8.84	88.73	-3.95	39.57	64.38
MX24	0	25	75	-3.75	41.25	62.50	0.92	32.31	66.76
MX25	33	33	33	-28.53	131.82	-3.29	30.90	14.63	54.47
MX26	33	33	33	41.44	-2.8	61.36	20.28	39.39	40.33
MX27	33	33	33	3.72	63.63	32.65	27.32	16.63	56.05
MX28	33	33	33	32.45	7.71	59.84	37.62	-5.74	68.13
MX29	67	33	0	18.34	123.67	-42.01	58.41	47.98	-6.39
MX30	33	67	0	7.15	110.08	-17.23	26.78	72.59	0.63
MX31	67	0	33	61.87	30.08	8.05	58.73	37.46	3.81
MX32	33	0	67	-9.42	88.44	20.98	37.37	-4.67	67.30
MX33	0	67	33	-19.17	117.86	1.31	-13.57	110.83	2.74
MX34	0	33	67	25.67	18.42	55.91	-10.97	94.05	16.92
MX35	100	0	0	64.31	61.06	-25.37	95.28	3.29	1.43
MX36	0	100	0	-26.74	101.15	25.59	-18.82	89.23	29.59
MX37	0	100	0	-11.89	146.86	-34.98	-15.31	157.66	-42.35
MX38	0	100	0	-42.68	132.74	9.94	-31.82	117.37	14.45
MX39	0	100	0	67.82	19.54	12.63	17.30	114.64	-31.95
MX40	0	100	0	-0.80	89.30	11.50	-9.69	109.86	-0.17
MX41	0	0	100	14.33	-31.69	117.36	-39.95	81.83	58.12
MX42	0	0	100	-12.41	99.39	43.02	8.07	1.38	90.55

rms difference between known and calculated compositions) for unmixing of data set E is 0.404. If the value 0.33 is assigned to all calculated proportions, the rms error values would be 0.347. This case therefore exhibits no skill in determining image composition by ice type.

The samples of data set E are 128×128 pixels in size compared to 64×64 for data set B where the unmixing accuracy was calculated to be 0.632. Based on sample size and random noise cancellation, one would expect the error in data set E to be one-half of the value for data set B, or 0.316, rather than the 0.404 measured here. The departure of the accuracy for these larger samples from the random-noise-based predictions may be due to several factors, including seam effects in the synthetic mixtures, different relative weightings of first-year, second-year, and multiyear ice types in data sets B and E, and the small statistical sampling represented by 42 data points.

2. Texture Plus MEAN

The 10 texture measures included in data set C were combined with mean image intensity, MEAN, to construct an 11-variable data set called F (see Appendix H). The data from the 42 mixture samples and the class means from data set A were, as before, SIFTed and processed by CABFAC to accomplish a transformation into a three-dimensional reduced feature space. The calculated

transformation matrix in this case is

$$TV = \begin{bmatrix} -0.050 & 0.429 & -0.091 \\ 0.441 & 0.032 & 0.208 \\ -0.037 & 0.449 & -0.184 \\ 0.407 & 0.000 & 0.182 \\ 0.002 & 0.457 & 0.037 \\ 0.454 & 0.020 & -0.078 \\ -0.044 & 0.383 & -0.099 \\ 0.455 & 0.063 & -0.020 \\ -0.020 & 0.414 & -0.235 \\ 0.470 & 0.025 & -0.289 \\ -0.007 & 0.286 & 0.851 \end{bmatrix} . \quad (VII - 3)$$

Class mean end members for this case are

$$\begin{aligned} \vec{E}_{FY} &= [\ 0.922 \quad 0.082 \quad -0.003] \\ \vec{E}_{SY} &= [\ 0.443 \quad 0.357 \quad 0.200] \\ \vec{E}_{MY} &= [\ 0.182 \quad 0.788 \quad 0.036] . \end{aligned} \quad (VII - 4)$$

Figure VII-9 is a plot of data set F and the mixing polytope constructed from the end members of Eq. (VII-4). Ice-type compositions based on this polytope and the QMODEL code are given in Table VII-1 beside the the data set E values previously cited. The accuracy of the calculated mixing proportions is 0.211, which is approximately one-half of the value for texture only in data set E. We see here the same accuracy relationships between texture only and texture plus mean cases that was observed previously for pure sample data sets B and C. The accuracy for F is again higher than the value of 0.167, which is predicted from the results of the smaller sample data set C and random noise cancellation considerations.

C. UNMIXING WITH EXTREME SAMPLES AS END MEMBERS

Up to this point we have used class means calculated from data set A, which contains only pure samples, as end members for unmixing calculations.

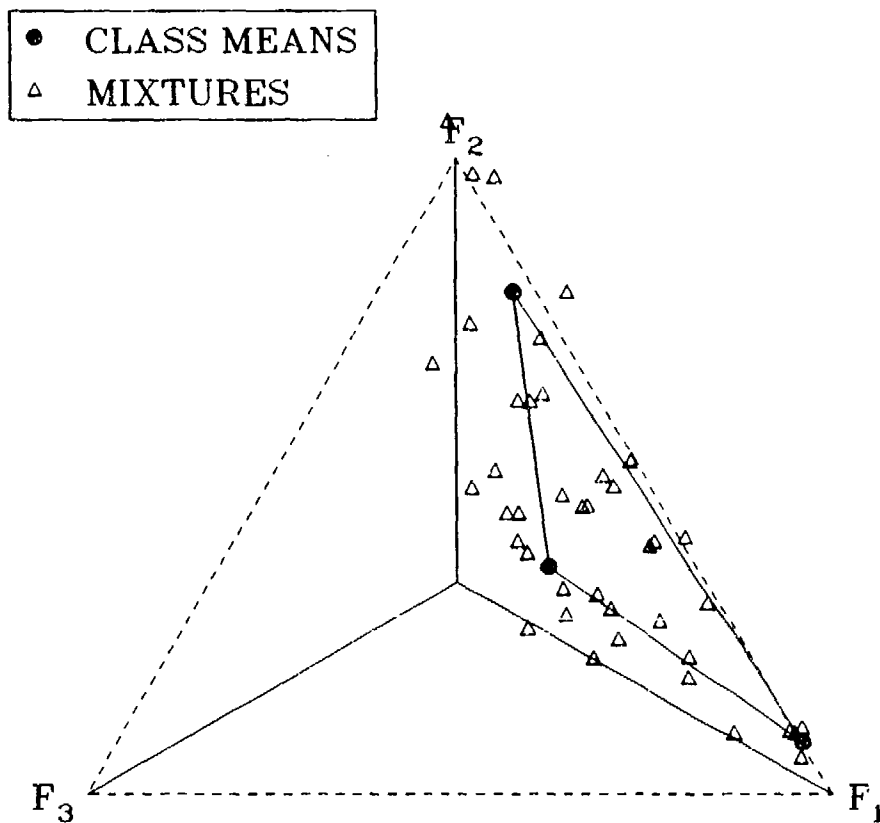


FIGURE VII-9. Plot of data set F in the reduced feature space. Solid dots indicate mean positions for samples belonging to FY, SY, and MY ice classes in data set A.

This choice of end members is reasonable, since one could assume that the mean vector for a number of pure samples of a single ice type would be the "best" representation of a sample consisting of 100% of that ice type. However, mean vectors, which intuitively seem appropriate or even optimal as end members, may not in fact be the best choice when best is judged in terms of the resulting accuracy of calculated sample proportions. Furthermore, one may not always have a set of pure samples of known composition from which to calculate means. Therefore, as a practical matter, alternate techniques for finding end members may be necessary or even preferable in some cases. Finding end-member vectors is the major function of QMODEL. Several options are offered by that computer code.

The simplest approach would be to use the axes of the varimax space as end member vectors. Certainly this approach satisfies the linear model (Eq. (VI-1)) exactly, where the \vec{E}_k vectors are the varimax axes and the a_{ik} coefficients are the coordinates of the data points in varimax space. This approach has several disadvantages, however. First, the coordinates of the end members in the original feature space may not always be positive. In other words, end members coincide with physically unrealizable sample vectors so that a pure sample of a given end member could never exist. A second disadvantage is that the constraint of maintaining end members on the edge of the positive orthant is overly restrictive. Data sets, such as the texture data studied here, where we have a complete data cloud formed by samples that are known to be mixtures of known end members, show end member locations (i.e., vertices of the data cloud) which fall well away from the varimax feature space axes. Clearly, this option is inappropriate for the present image texture data, so the possibility of F_1 , F_2 , and F_3 as end-member vectors was not pursued in this study.

We move directly to the approach of Imbrie and Van Andel (1964), which was to consider the end members to be unit vectors that pass through the positions of real sample vectors occupying mutually extreme locations in the reduced feature space. This approach, of course, assumes that all end members are represented in the data set. It further assumes that no significant outliers result from erroneous samples. A set of extremal samples can be determined by an iterative method described by Manson and Imbrie (1964). Once the end-member samples are identified, the compositions of each sample, a_{ik} , can be derived using the oblique projection method given by Imbrie (1963). These techniques for locating and using extreme samples as end members are offered as options in the QMODEL code. Another option is to find extremal samples in the original feature space. This second type of extremal end member was not investigated. End members taken as the most mutually extreme samples in varimax space are sometimes call Imbrie oblique end members. The terminology "Imbrie end members" will be used here.

1. Texture Measures Only

Data set E was unimixed using the three extreme samples (MX9, MX38, and MX1 for FY, SY, and MY ice, respectively) identified by QMODEL as being the Imbrie end members. Figure VII-10 shows data set E in reduced feature space with a polytope formed by connecting these end members. End-member vectors coinciding with MX9, MX38, and MX1 are given below.

$$\vec{E}_{FY} = [0.912 \quad 0.093 \quad -0.005]$$

$$\vec{E}_{SY} = [0.290 \quad 0.381 \quad 0.329] \quad (VII - 5)$$

$$\vec{E}_{MY} = [0.080 \quad 0.943 \quad -0.023]$$

Table VII-2 gives the calculated ice-type proportions for each synthetic mixture sample. The rms composition error in this case was 0.289. This improvement is

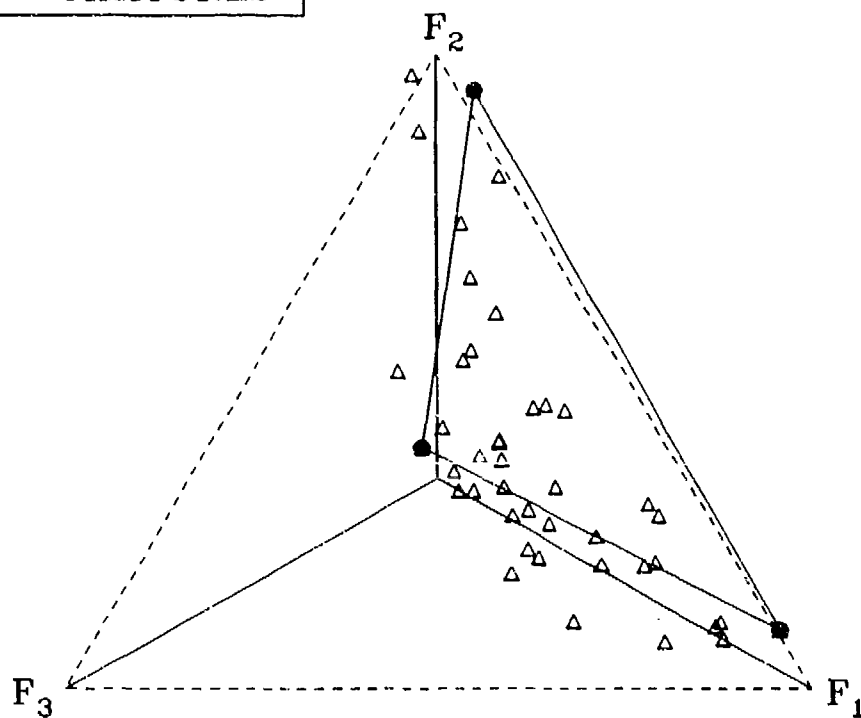


FIGURE VII-10. Plot of data set E in the reduced feature space. Solid dots indicate Imbrie end members.

TABLE VII-2. UNMIXING RESULTS FOR DATA SETS E AND F
Imbrie End Members

Sample	Known Proportions (%)			Set E - Texture Only			Set F - Texture + Mean		
	FY	SY	MY	FY	SY	MY	FY	SY	MY
MX1	0	0	100	0.00	0.00	100.00	11.86	-3.98	92.12
MX2	0	0	100	-15.12	16.01	99.12	0.00	0.00	100.00
MX3	0	0	100	-11.87	24.12	87.75	5.43	1.81	92.76
MX4	0	100	0	10.61	92.21	-2.82	-1.47	70.63	30.84
MX5	0	100	0	40.22	63.52	-3.74	18.47	66.21	15.32
MX6	0	100	0	52.85	47.16	-0.02	18.62	73.18	8.20
MX7	100	0	0	87.48	21.81	-9.29	98.24	3.72	-1.96
MX8	100	0	0	88.00	19.00	-7.00	100.00	0.00	0.00
MX9	100	0	0	100.00	0.00	0.00	101.02	6.81	-7.84
MX10	50	25	25	29.93	77.67	-7.60	52.88	14.05	33.07
MX11	50	25	25	39.28	51.11	9.61	51.21	16.88	31.91
MX12	25	50	25	56.40	52.27	-8.67	38.64	55.21	6.14
MX13	25	50	25	21.78	62.82	15.40	21.94	39.64	38.42
MX14	25	25	50	23.55	66.85	9.60	29.19	33.92	36.89
MX15	25	25	50	9.49	52.44	38.07	18.85	22.87	58.28
MX16	50	0	50	33.71	69.33	-3.04	62.77	1.65	35.58
MX17	0	50	50	8.84	7.05	84.11	-3.39	33.70	69.69
MX18	50	50	0	65.11	18.88	16.01	54.23	30.99	14.77
MX19	25	75	0	36.38	80.57	-16.95	31.96	51.59	16.45
MX20	75	25	0	90.32	21.53	-11.85	79.46	26.96	-8.41
MX21	75	0	25	61.82	20.25	17.94	69.82	7.61	22.57
MX22	25	0	75	5.87	36.40	57.72	30.06	-5.71	75.65
MX23	0	75	25	14.19	34.42	51.39	9.73	33.85	56.42
MX24	0	25	75	7.61	58.10	34.29	13.99	29.43	56.58
MX25	33	33	33	13.15	95.58	-8.73	40.24	17.07	42.69
MX26	33	33	33	36.39	30.25	33.36	30.97	31.66	37.37
MX27	33	33	33	21.89	63.41	14.69	37.11	18.49	44.41
MX28	33	33	33	31.23	36.37	32.40	43.09	5.25	48.66
MX29	67	33	0	55.06	79.24	-24.29	64.42	33.10	2.48
MX30	33	67	0	40.07	77.95	-18.01	36.74	49.42	13.84
MX31	67	0	33	66.36	35.30	-1.67	64.68	27.26	8.05
MX32	33	0	67	17.38	75.52	7.10	45.87	5.86	46.27
MX33	0	67	33	17.55	88.20	-5.75	1.43	74.01	24.56
MX34	0	33	67	28.23	41.92	29.85	3.68	64.52	31.80
MX35	100	0	0	78.63	44.95	-23.58	96.66	5.25	1.91
MX36	0	100	0	5.00	84.91	10.19	-3.21	62.54	40.67
MX37	0	100	0	33.79	95.77	-29.56	0.00	100.00	0.00
MX38	0	100	0	0.00	100.00	0.00	-14.56	79.20	35.36
MX39	0	100	0	68.64	30.05	1.31	28.51	73.44	-1.95
MX40	0	100	0	25.84	73.31	0.85	4.83	73.14	22.03
MX41	0	0	100	1.37	28.46	70.17	-21.75	60.28	61.46
MX42	0	0	100	-10.45	88.77	21.68	20.19	11.74	68.07

significant over the rms error of 0.404 that resulted from using mean vectors as end members.

2. Texture Plus MEAN

Data set F was also unmixed using Imbrie end members. In this case samples MX8, MX37, and MX2 were selected by QMODEL.

$$\begin{aligned}\vec{E}_{FY} &= [0.909 \quad 0.105 \quad -0.014] \\ \vec{E}_{SY} &= [0.462 \quad 0.262 \quad 0.276] \\ \vec{E}_{MY} &= [-0.040 \quad 1.059 \quad -0.019]\end{aligned}\tag{VII-6}$$

Figure VII-11 is a plot of this polytope. Table VII-2 gives the calculated sample compositions, which result in an rms composition error of 0.165. Again, this error is significantly smaller than the 0.211 reported when mean vectors were utilized as end members.

D. UNMIXING WITH DENEG END MEMBERS

Full *et al.* (1981) introduce another extremal method for locating end members. Whereas Imbrie (1964) used extremal points as vertices of the mixing polytope, Full *et al.* (1981) developed the DENEQ procedure, which uses extremal samples to position the sides of the polytope rather than the vertices. The end members are taken to be the intersections of the various sides thus determined. Under this approach, end members do not coincide with real samples. The ability to select end members well outside of the data cloud is the advantage (and also the disadvantage, as we will see later) of DENEG. Pure end members do not need to be present in the data set. Sides are projected to intersect at the presumed locations of missing end members.

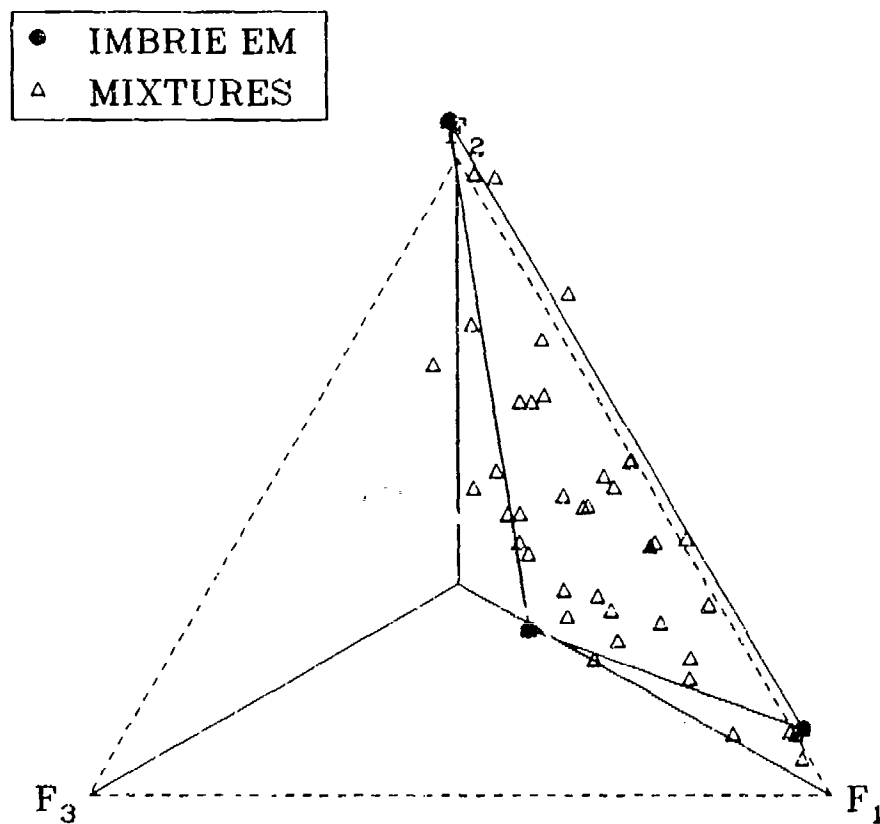


FIGURE VII-11. Plot of data set P in the reduced feature space. Solid dots indicate Imbrie end members.

One requirement of DENEG is that all sides of the polytope must be represented. This condition is true if each end member is missing from at least one sample. Another requirement for DENEG is that an initial polytope be available which approximately captures the shape of the data cloud. The DENEG algorithm starts with the initial polytope and iteratively moves the sides outward until the data cloud is enclosed, or until certain termination conditions are met. The steps of the algorithm are given in Full *et al.* (1981). Full *et al.* (1982) introduce FUZZY QMODEL, which incorporates clustering techniques from the field of pattern recognition to establish the polytope for DENEG initialization. Specifically, Full *et al.* (1982) used fuzzy clustering (Bezdek, 1981) to generate "cluster centers" within the data cloud. The cluster centers are taken as the initial end members, or vertices of the polytope, which DENEG subsequently expands to encompass the data. In this context, the advantageous characteristic of fuzzy clustering is that the shape of a polygon, formed by connecting the cluster centers, often assumes the shape of the data cloud, thereby satisfying one of the DENEG requirements for polytope initialization.

The reader is again referred to Full *et al.* (1981) for a complete description of the steps of the DENEG algorithm. Here, it will suffice to give only brief descriptions of the five "tuning" parameters associated with DENEG.

t_1 - The range of negative oblique space coordinate values that is considered negligible is from 0.0 to t_1 . Default value for t_1 is -0.05.

t_2 - The value of negative oblique space coordinates beyond which samples are considered to be "outliers," which will be ignored in DENEG calculations. Default values for t_2 is -0.25.

t_3 - All vertex coordinates in the original feature space must be larger than

t_3 . This parameter prohibits construction of end members that would be physically unrealistic in the original feature space. Default value for t_3 is -0.05.

t_4 - If the varimax coordinates of the new end members fall within a distance, t_4 , of the end members from the previous iteration, the DENEG procedure is terminated. The default value for t_4 is 0.05.

t_5 - The maximum number of DENEG iterations. The default value for t_5 is 10.

1. Texture Measures Only

FUZZY QMODEL was applied to data set E. Default values for t_1, t_3, t_4 , and t_5 were used, but t_2 was adjusted to vary the degree to which DENEG could expand the polytope outward from the fuzzy cluster centers. For each value of t_2 the end members were constructed, the unmixing was performed, and the rms difference between known and calculated proportions was computed. Figure VII-12 is a plot of the rms proportion error as a function of t_2 (called the DENEG cut-off value). For reference purposes the rms error for class means and for Imbrie end members are also plotted in the figure. Note in Figure VII-12 that for $t_2 = 0$ (i.e., fuzzy cluster centers are used as end members), the error is approximately 0.45. As t_2 increases and allows more polytope expansion by DENEG, the error drops until $t_2 = 0.4$; beyond which no further reduction in error occurs. This leveling off beyond 0.4 occurs because at that point, all 42 samples are included in the DENEG calculation so that any increase in t_2 has no effect. The mixing polytopes for $t_2 = -0.25$ and $t_2 = -0.4$ are shown in Figures VII-13 and Figure VII-14, respectively. Calculated composition values for the $t_2 = 0.4$ case are listed in Table VII-3. Note that in Figure VII-12

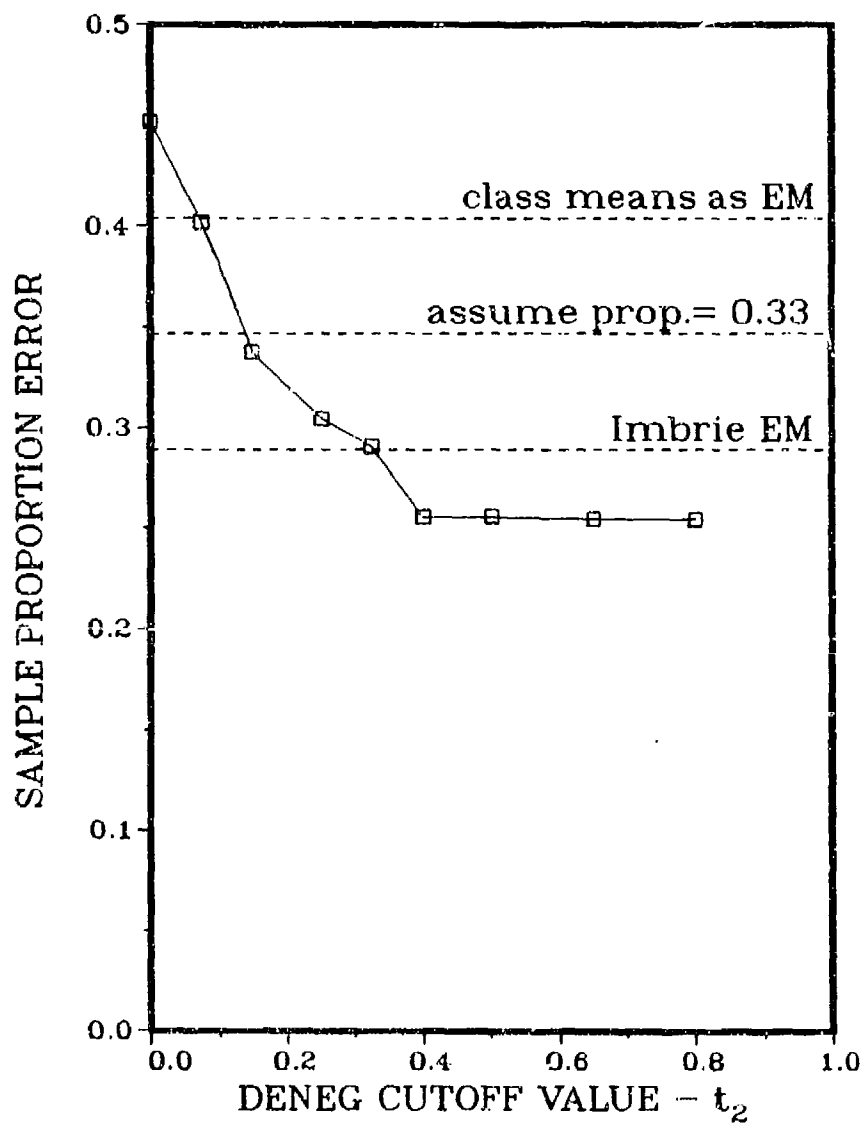


FIGURE VII-12. Error in calculated sample proportions as a function of the DENEG cut-off value for data set E. Various error reference levels are shown by dotted lines.

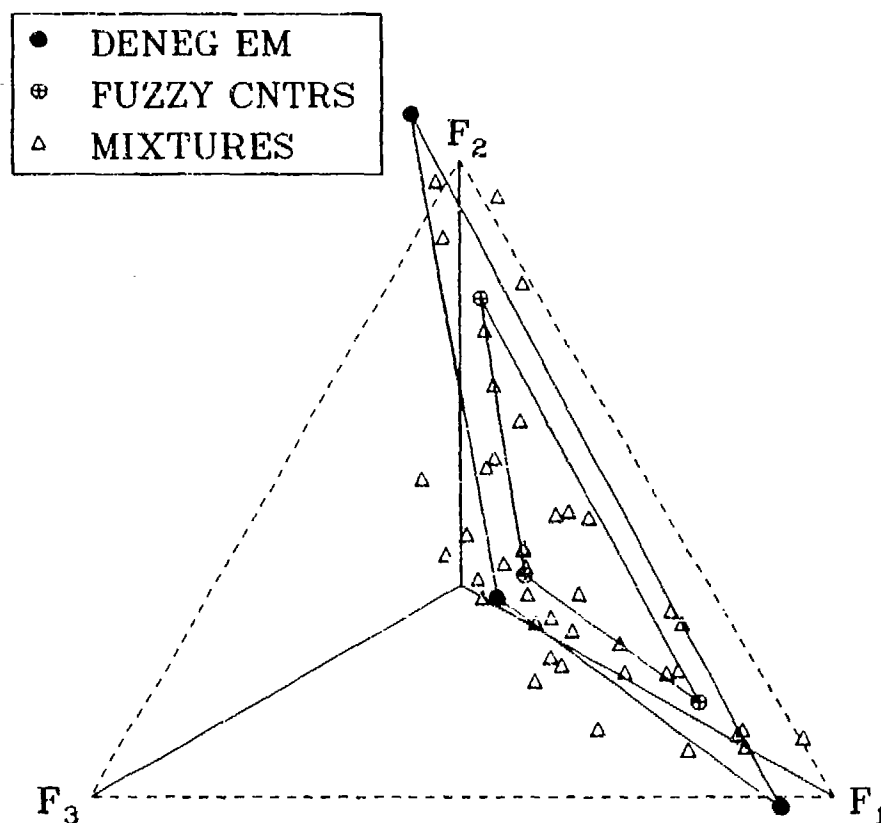


FIGURE VII-13. Plot of data set E in the reduced feature space. Solid dots indicate DENEG end members. Crossed dots indicate fuzzy cluster centers. $t_2 = -0.25$.

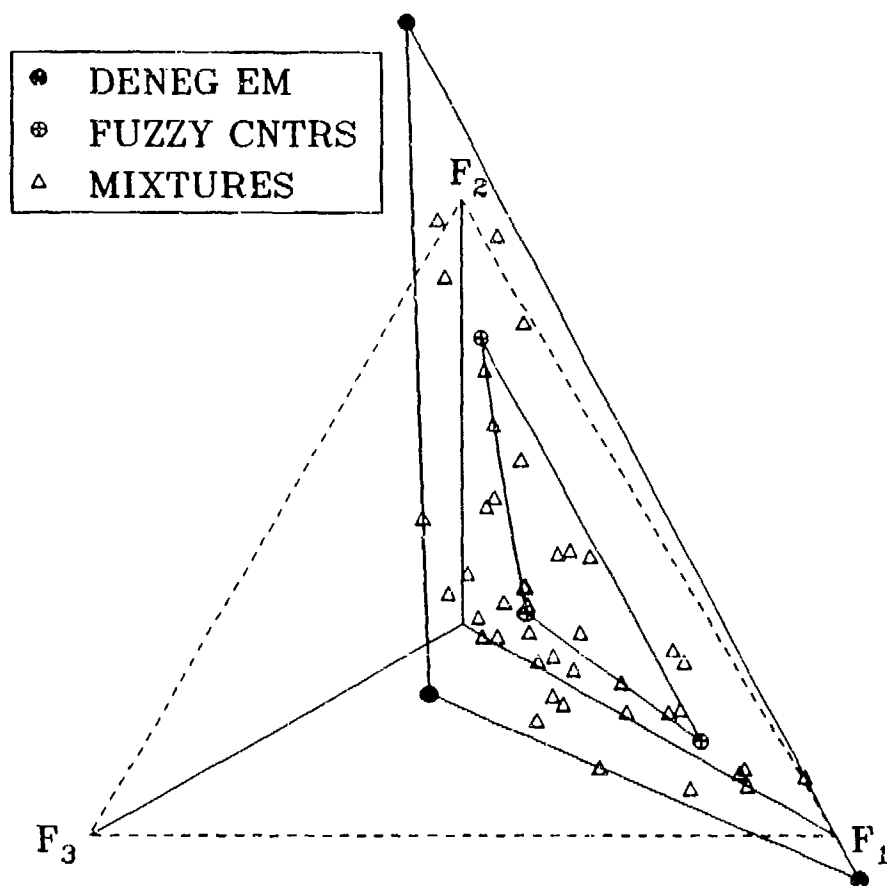


FIGURE VII-14. Plot of data set E in the reduced feature space. Solid dots indicate DENEG end members. Crossed dots indicate fuzzy cluster centers. $t_2 = -0.4$.

**TABLE VII-3. UNMIXING RESULTS FOR DATA SETS E AND F
DENEG End Members**

Sample	Known Proportions (%)			Set E - Texture Only			Set F - Texture + Mean		
	FY	SY	MY	FY	SY	MY	FY	SY	MY
MX1	0	0	100	16.88	4.18	78.95	20.46	1.19	78.35
MX2	0	0	100	4.78	15.37	79.85	12.23	4.19	83.58
MX3	0	0	100	6.12	20.59	73.30	16.19	5.66	78.15
MX4	0	100	0	14.02	64.58	21.40	15.01	59.39	25.61
MX5	0	100	0	37.41	44.41	18.17	28.97	56.11	14.92
MX6	0	100	0	47.85	33.09	19.06	29.46	61.56	8.98
MX7	100	0	0	74.30	14.92	10.79	82.42	7.98	9.60
MX8	100	0	0	74.97	13.04	11.99	83.47	5.09	11.44
MX9	100	0	0	85.25	0.00	14.75	84.56	10.43	5.01
MX10	50	25	25	28.83	54.18	17.00	50.66	15.65	33.69
MX11	50	25	25	38.14	36.26	25.60	49.63	17.85	32.53
MX12	25	50	25	49.71	36.31	13.98	42.75	47.70	9.56
MX13	25	50	25	24.88	44.72	30.40	30.01	35.37	34.62
MX14	25	25	50	25.65	47.31	27.04	34.86	30.97	34.17
MX15	25	25	50	17.62	38.39	44.00	26.89	22.24	50.87
MX16	50	0	50	32.33	48.51	19.16	57.03	6.05	36.92
MX17	0	50	50	22.15	8.46	69.39	11.64	30.51	57.86
MX18	50	50	0	59.33	13.92	26.76	52.54	28.90	18.55
MX19	25	75	0	32.92	55.82	11.26	37.79	44.80	17.41
MX20	75	25	0	76.27	14.61	9.11	70.40	27.54	2.06
MX21	75	0	25	56.93	14.95	28.12	62.38	10.77	26.85
MX22	25	0	75	16.90	27.95	55.15	33.33	0.00	66.67
MX23	0	75	25	22.81	26.30	50.89	20.99	30.74	48.27
MX24	0	25	75	15.71	42.20	42.09	23.79	27.32	48.89
MX25	33	33	33	15.39	66.70	17.91	41.82	17.90	40.28
MX26	33	33	33	38.45	22.62	38.94	36.01	29.21	34.78
MX27	33	33	33	24.90	45.10	30.00	39.66	18.98	41.36
MX28	33	33	33	34.25	26.87	38.88	45.34	8.71	45.95
MX29	67	33	0	45.84	54.16	0.00	59.91	30.64	9.45
MX30	33	67	0	35.73	53.94	10.34	41.08	43.15	15.77
MX31	67	0	33	58.38	24.70	16.92	59.79	26.08	14.13
MX32	33	0	67	20.48	53.29	26.23	45.22	9.19	45.59
MX33	0	67	33	19.21	61.64	19.15	17.26	62.05	20.69
MX34	0	33	67	31.58	30.66	37.76	18.35	54.66	27.00
MX35	100	0	0	65.71	30.56	3.73	81.37	9.16	9.47
MX36	0	100	0	11.00	59.93	29.07	13.33	53.05	33.62
MX37	0	100	0	29.48	65.96	4.56	17.65	82.35	0.00
MX38	0	100	0	5.91	70.16	23.92	6.15	65.96	27.89
MX39	0	100	0	60.51	21.14	18.35	36.52	61.85	1.63
MX40	0	100	0	26.51	51.47	22.02	19.63	61.40	18.97
MX41	0	0	100	14.69	22.90	62.41	0.00	51.12	48.88
MX42	0	0	100	0.00	63.19	33.81	27.25	13.55	59.21

the DENEg error line goes below the rms error level for Imbrie end members. DENEg-derived end members are, therefore, slightly superior to the others for this particular set of synthetic mixture image texture data. End members for $t_2 = -0.4$ for the texture only case are

$$\begin{aligned}\vec{E}_{FY} &= [1.065 \quad -0.069 \quad 0.004] \\ \vec{E}_{SY} &= [0.343 \quad 0.223 \quad 0.434] \quad . \\ \vec{E}_{MY} &= [-0.210 \quad 1.279 \quad -0.069]\end{aligned}\quad (VII - 7)$$

2. Texture Plus MEAN

Data set F was also processed by FUZZY QMODEL using a range of values for the DENEg parameter t_2 . The rms error values for this data set are plotted in Figure VII-15. As previously in Figure VII-12, the error values in Figure VII-15 decrease as the polytope is expanded. When all samples are included, the error value becomes fixed. The difference between this case and the texture-only case is that for texture plus MEAN, the DENEg error never reaches the value for Imbrie end members. The polytope calculated by DENEg for $t_2 = 0.325$ is shown in Figure VII-16. The reason for the poor performance of DENEg as compared to extreme samples is apparent from Figure VII-16. DENEg has projected an end member out of the positive orthant by a significant distance. The possibility of this happening is one disadvantage of DENEg. Polytope sides based on extreme samples may intersect at strange places. Even more extreme examples of this problem will be seen later.

The end-member vectors from DENEg with $t_2 = 0.325$ for the texture

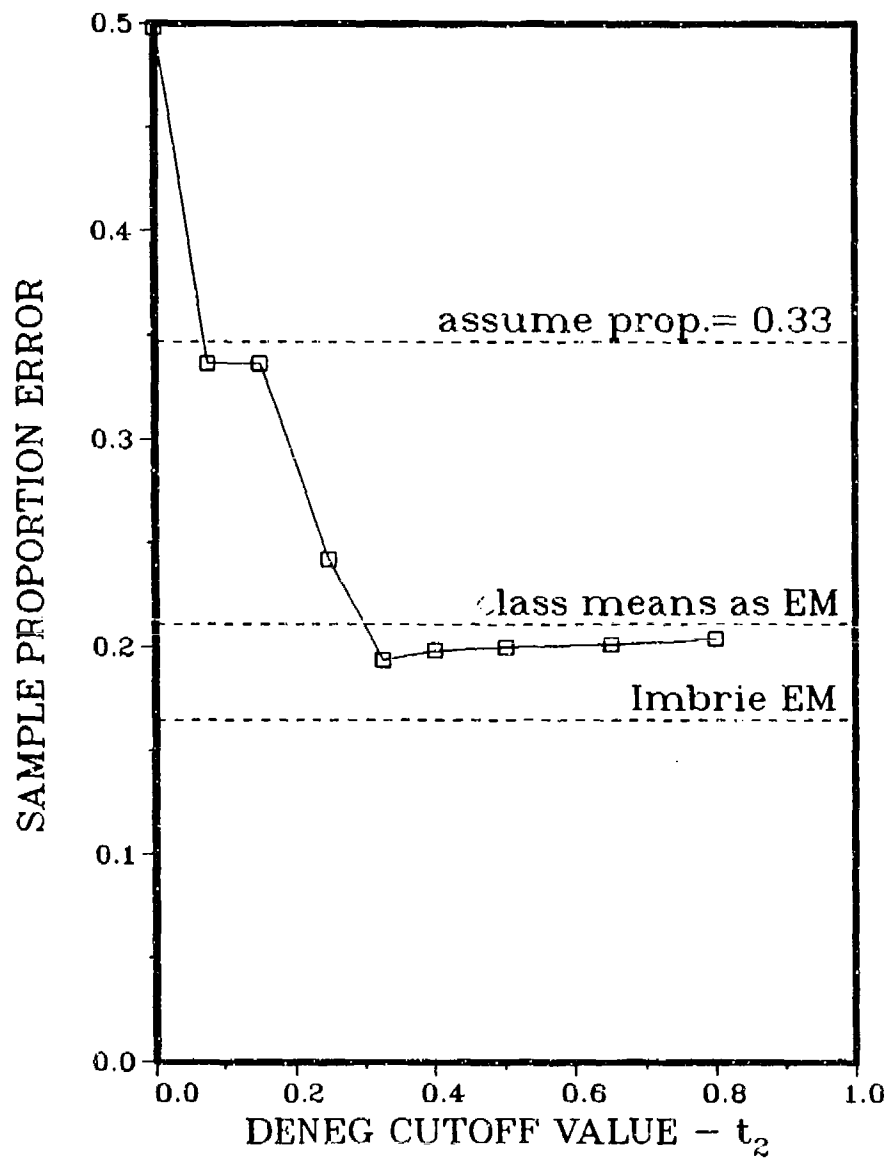


FIGURE VII-15. Error in calculated sample proportions as a function of the DENEG cut-off value for data set F. Various error reference levels are shown by dotted lines.

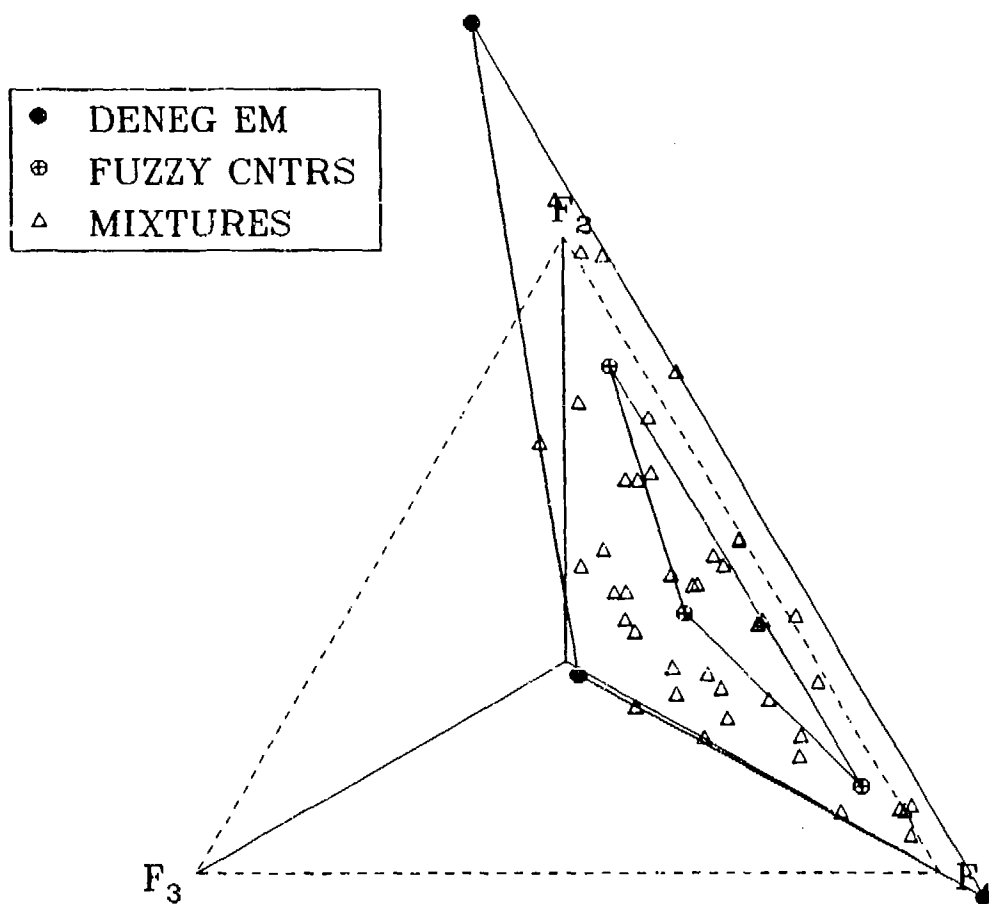


FIGURE VII-16. Plot of data set F in the reduced feature space. Solid dots indicate DENEG end members. Crossed dots indicate fuzzy cluster centers. $t_2 = -0.325$.

plus mean case are

$$\begin{aligned}\vec{E}_{FY} &= [1.077 \quad -0.039 \quad -0.038] \\ \vec{E}_{SY} &= [0.360 \quad 0.311 \quad 0.328] \\ \vec{E}_{MY} &= [-0.219 \quad 1.336 \quad 0.045]\end{aligned}\quad (VII-8)$$

Calculated proportion values based on these end members are given in the last three columns of Table VII-3.

E. SUMMARY OF ANALYSIS OF SYNTHETIC MIXTURES

Unmixing calculations have been performed on the synthetic mixture images using end members generated by several methods. Based on texture measures only, image composition was calculated to an rms proportion error of 0.254 in the best case. The best case for texture plus MEAN was an rms error of 0.165. These errors extrapolate to 0.08 and 0.05, respectively, for full-frame KRMS images. The analysis of synthetic mixtures has suggested that natural KRMS images can be unmixed with useful accuracy, provided the problem of unequal mean image intensities for different ice classes can be shown to be solved satisfactorily. Section VIII will examine the unequal means problem.

This synthetic data set revealed several things about selection of end members. First, class means as end members have been shown to be the least suitable in terms of accuracy of the resulting calculated proportions. This result is somewhat surprising, since the mean vector for a given class of ice should be the best possible estimate of what a pure sample would look like in feature space. The other two methods of obtaining end members - one using extremal samples as end members, the other using extremal samples to define polytope sides - performed better than the class means as end members. The performance of these extremal methods is also surprising because of the high noise levels in

the data. One would expect that relatively noise-free data would be required to obtain useful results from methods that key on extreme samples.

DENEG (the side extreme method) and Imbries end members (the vertex extreme method) gave comparable results. The relative performance of these two changed with data set. So far in this study there seems to be no clearly superior method. Vertex extremes appear to be more stable, having less chance of wildly divergent solutions. DENEG, however, can excel if the data cloud supports an orderly convergence to a reasonable solution, which is not guaranteed. The rule-of-thumb the author has adopted at this point is to use Imbries end members if all end members are known to be represented in the data set. If not all end members are present, DENEG is the only choice, but one should be aware of the possibility that unsatisfactory end members are possible; caution is advised.

VIII. UNMIXING WITH VARIABLE MEANS

In Section III several desirable characteristics of texture measures were identified. One of the most important characteristics to linear unmixing is obviously a linear behavior in the formation of mixtures. In Section IV we found that only 6 of the 25 candidate texture measures mix linearly. We found further that linear behavior could be attributed to 12 texture measures if the mean intensity levels in the components of the mixtures were identical. Our synthetic mixtures, therefore, were forced to uniform mean intensities so that we could take full advantage of all 12 texture measures. Now as we look toward natural mixtures that will not have invariant mean intensities for all ice types, we must reduce the suite of texture measures from 12 to 6, and also begin to analyze mixtures that do not have equal means. This section addresses these questions. First, the synthetic mixtures from the previous section will be unmixed using only the 6 more robust measures that handle mean shifts. The objective is to see how much ice-type discrimination power is lost by discarding half of the variables previously incorporated in the analysis. Second, the 42 synthetic mixtures will be recreated without any mean adjustment, and the analysis repeated to measure the degradation in unmixing performance that might result from nonadjusted means. If both the reduction in variables and introduction of mean shifts into the synthetic mixture have relatively minor impact on unmixing accuracy, then we will be ready to go on to natural mixtures in Section IX.

A. REDUCTION TO SIX VARIABLES

Section IV identified MEAN, HOM3, HOM15, LM2, LM3, LM4, and LM5 (6 texture variables plus MEAN) as the only variables among the 25 evaluated that would mix linearly in the most general case, where various parts of the mix-

ture would have textures superimposed on varying overall background brightness levels. The synthetic mixtures of data set D were reanalyzed using only these variables. The 6 texture variables were extracted from D and SIFTed to form data set G listed in Appendix I. The four texture measures HOM3, HOM15, LM2, and LM3 were combined with MEAN to form the five-dimensional SIFTed data set, called H, listed in Appendix J.

1. Texture Measures Only

The six-dimensional data set G was processed by CABFAC in the same manner described in association with previous data sets. The resulting transformation matrix that will put G into a three-dimensional, reduced varimax feature space is

$$TV = \begin{bmatrix} 0.565 & -0.045 & 0.154 \\ -0.266 & 0.578 & 0.498 \\ 0.540 & -0.036 & 0.255 \\ -0.120 & 0.571 & 0.096 \\ 0.239 & -0.080 & 0.654 \\ 0.497 & 0.575 & -0.476 \end{bmatrix}. \quad (VIII - 1)$$

Figure VIII-1 is a plot of G in the reduced feature space. Figure VIII-2 shows several unmixing polytopes calculated by QMODEL. The end-member coordinates for the polytopes shown in Figure VIII-2 are given below.

Class means from Section VI:

$$\begin{aligned} \vec{E}_{FY} &= [0.634 \quad 0.133 \quad 0.233] \\ \vec{E}_{SY} &= [0.350 \quad 0.201 \quad 0.450] \\ \vec{E}_{MY} &= [0.148 \quad 0.579 \quad 0.273] \end{aligned} \quad (VIII - 2)$$

Imbrie end members:

$$\begin{aligned} \vec{E}_{FY} &= [0.664 \quad 0.154 \quad 0.182] \\ \vec{E}_{SY} &= [0.288 \quad 0.176 \quad 0.536] \\ \vec{E}_{MY} &= [0.093 \quad 0.855 \quad 0.051] \end{aligned} \quad (VIII - 3)$$

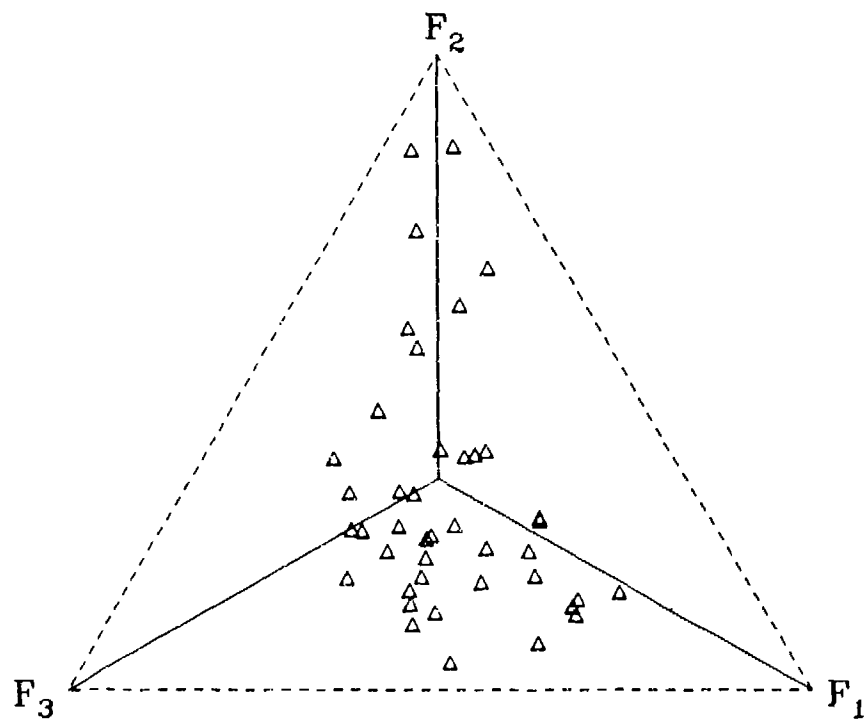
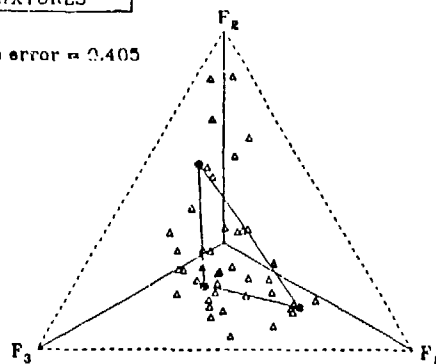


FIGURE VIII--1. Plots of data set G in the reduced feature space.

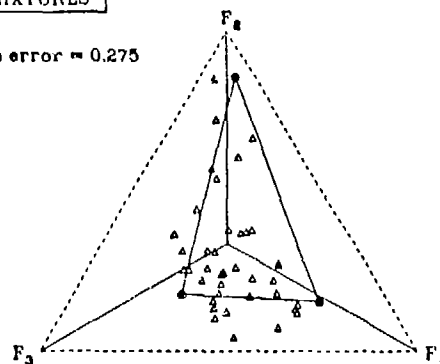
• CLASS MEANS
▲ MIXTURES

rms error = 0.405



• IMBRIE EM
▲ MIXTURES

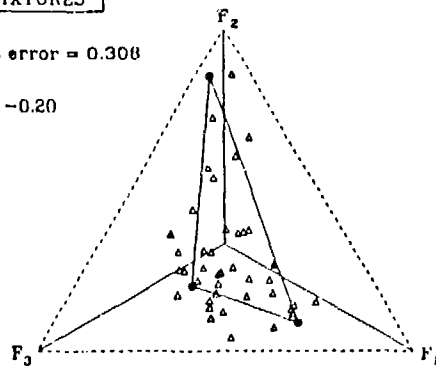
rms error = 0.275



• DENEG EM
▲ MIXTURES

rms error = 0.308

$t_2 = -0.20$



• DENEG EM
▲ MIXTURES

rms error = 0.288

$t_2 = -0.40$

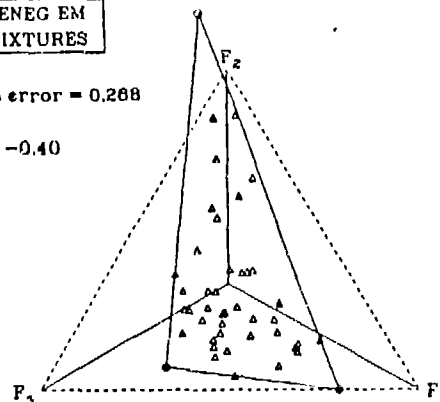


FIGURE VIII-2. Plots of data set G in the reduced feature space with various superimposed mixing polytopes.

DENEG end members with $t_2 = -0.2$:

$$\begin{aligned}\vec{E}_{FY} &= [0.921 \quad 0.126 \quad 0.368] \\ \vec{E}_{SY} &= [0.516 \quad 0.332 \quad 0.790] . \\ \vec{E}_{MY} &= [0.047 \quad 0.991 \quad 0.129]\end{aligned}\tag{VIII - 4}$$

DENEG end members with $t_2 = -0.4$:

$$\begin{aligned}\vec{E}_{FY} &= [0.966 \quad -0.001 \quad 0.260] \\ \vec{E}_{SY} &= [0.424 \quad 0.101 \quad 0.900] . \\ \vec{E}_{MY} &= [-0.135 \quad 0.991 \quad -0.018]\end{aligned}\tag{VIII - 5}$$

Note that the DENEG end members with $t_2 = -0.4$ give the best unmixing accuracy of 0.268, which compares favorably with the best texture-only error value of 0.254 from Section VII, where 12 rather than 6 texture variables were included. We conclude that the texture information lost by eliminating 6 texture variables is not significant.

2. Texture Plus MEAN

The five-dimensional texture plus MEAN data set was also processed by CABFAC to produce the reduced feature space data points shown in Figure VIII-3. The transformation to reduced varimax feature space in this case is

$$TV = \begin{bmatrix} 0.450 & -0.106 & 0.539 \\ -0.096 & 0.726 & 0.278 \\ 0.520 & -0.035 & 0.471 \\ -0.135 & 0.606 & 0.109 \\ 0.706 & 0.308 & -0.632 \end{bmatrix} .\tag{VIII - 6}$$

Figure VIII-4 shows several unmixing polytopes calculated by QMODEL for data set H. The end-member coordinates for the polytopes shown are as follows.

Class means from Section VI:

$$\begin{aligned}\vec{E}_{FY} &= [0.731 \quad 0.103 \quad 0.166] \\ \vec{E}_{SY} &= [0.311 \quad 0.271 \quad 0.418] . \\ \vec{E}_{MY} &= [0.171 \quad 0.643 \quad 0.186]\end{aligned}\tag{VIII - 7}$$

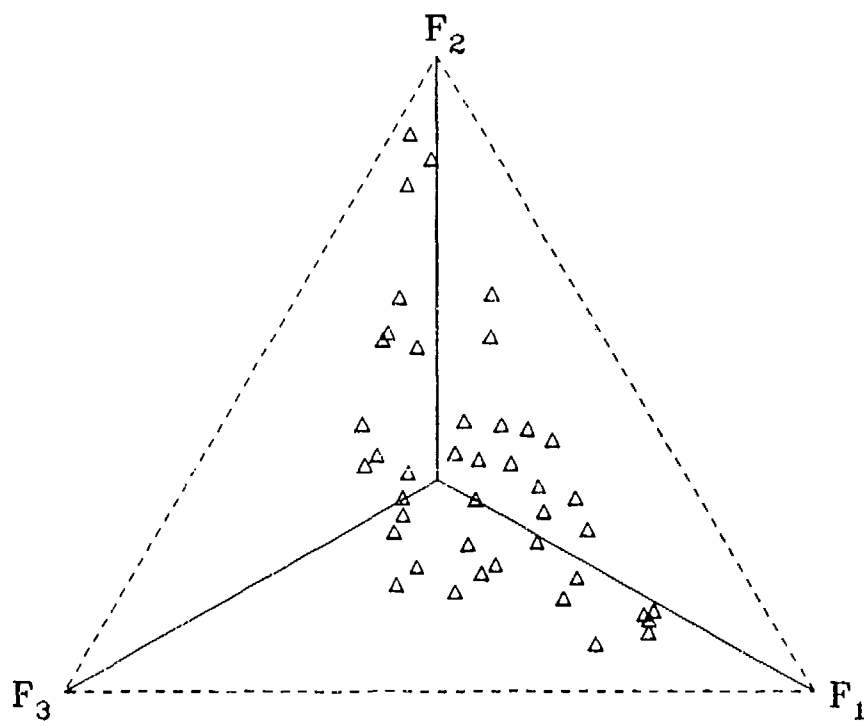


FIGURE VII-3. Plots of data set H in the reduced feature space.

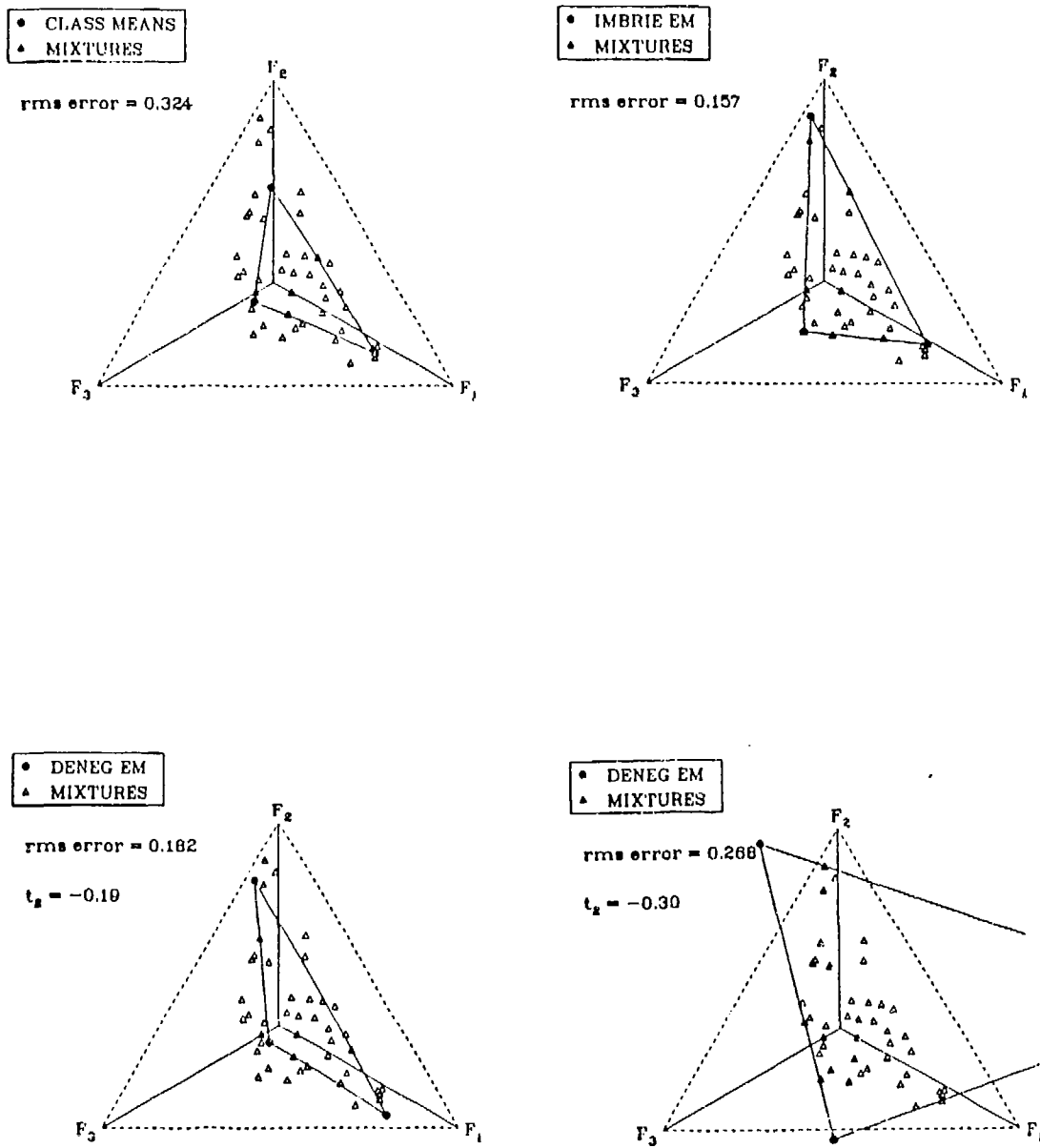


FIGURE VIII-4. Plots of data set H in the reduced feature space with various superimposed mixing polytopes.

Imbrie end members:

$$\begin{aligned}\vec{E}_{FY} &= [0.725 \quad 0.126 \quad 0.150] \\ \vec{E}_{SY} &= [0.360 \quad 0.167 \quad 0.473] \quad (VIII - 8) \\ \vec{E}_{MY} &= [0.025 \quad 0.877 \quad 0.099]\end{aligned}$$

DENEG end members with $t_2 = -0.19$:

$$\begin{aligned}\vec{E}_{FY} &= [0.973 \quad 0.051 \quad 0.224] \\ \vec{E}_{SY} &= [0.574 \quad 0.478 \quad 0.665] \quad (VIII - 9) \\ \vec{E}_{MY} &= [0.034 \quad 0.981 \quad 0.194]\end{aligned}$$

DENEG end members with $t_2 = -0.3$:

$$\begin{aligned}\vec{E}_{FY} &= [0.841 \quad 0.256 \quad -0.476] \\ \vec{E}_{SY} &= [0.676 \quad -0.045 \quad 0.736] \quad (VIII - 10) \\ \vec{E}_{MY} &= [-0.192 \quad 0.952 \quad 0.240]\end{aligned}$$

Note that the Imbrie end members give best unmixing accuracy of 0.157, which is slightly better than the best texture plus MEAN error value of 0.165 from Section VII, where 10 texture measures plus MEAN were included. Again the conclusion is that elimination of 6 texture measures does not significantly impact the accuracy of calculated image proportions.

Note in Figure VIII-4, when $t_2 = -0.3$, the DENEG solution loses the initial polytope shape and pushes an end member off the edge of the plot. This possibility of producing end members well away from the data cloud requires caution in using DENEG to select end members.

B. MIXTURES WITH MEAN SHIFTS

Synthetic mixture images MX1 through MX42, shown in Figures VII-1 through VII-7, were recreated by mosaicking pure samples as before, but this

time the adjustment of mean intensity to a value of 128 prior to mosaicking was not performed. These synthetic mixtures with variable mean intensities are called NX1 through NX42. Figure VIII-5 is a representative example of the NX series of mixture samples. The rest of the NX images are not shown here, since they are similar to Figures VII-1 through VII-7, except that the background brightness level is different for each component of the mixture. The 6 texture values calculated for the NX series of synthetic images are listed in Appendix K and are given the name data set I. The MEAN variable assumes the same value for the NX series of mixtures as it did in the MX series, since the previous MEAN values were calculated without the mean adjustment being performed. A SIFTed version of I is called J and is listed in Appendix L. Likewise, a SIFTed data set containing HOM3, HOM15, LM2, LM3, and MEAN is called K and is listed in Appendix M.

Table K-1 contains values of texture measures from synthetic mixtures with variable means. If Table K-1 is compared with the comparable values in Table F-1, where mean intensities were normalized, then one will find little change in the numerical values of the texture measures. Since the values of the raw data are only slightly altered by introducing shifts in the mean image intensity for each constituent of the mixture, then it is perhaps obvious that the NX series of synthetic mixtures can be unmixed with approximately the same accuracy as the MX mixtures. However, for the sake of completeness, limited unmixing results will be presented for the NX series of mixtures with variable means.

1. Texture Measures Only

Figure VIII-6 is a plot of data set J in the reduced varimax space. Imbrie end members and DENEG end members were applied to the unmixing of these



FIGURE VIII-5. Typical synthetic mixture image (NX2) without mean intensity level adjusted prior to mosaicking.

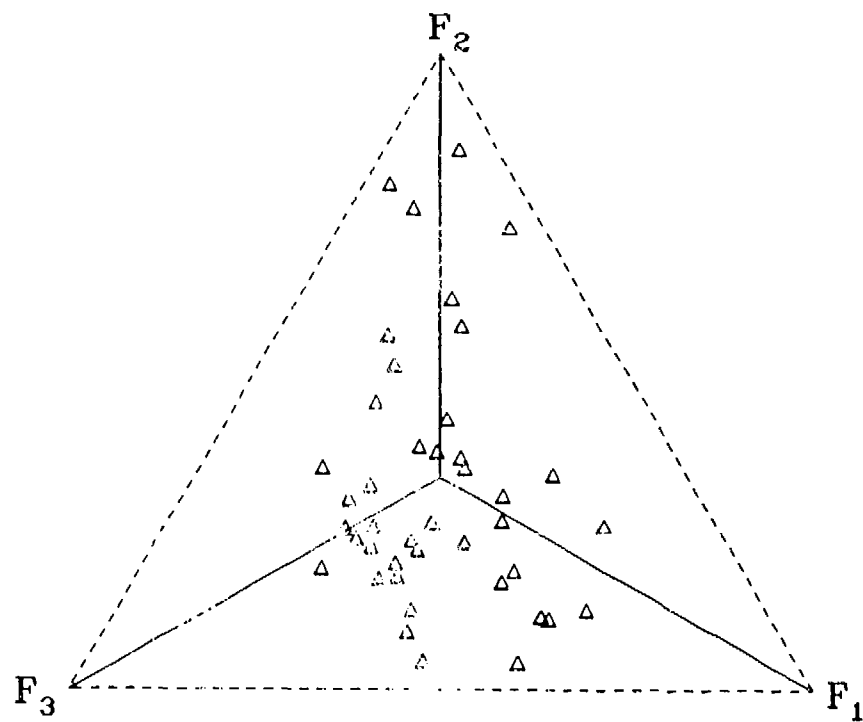


FIGURE VIII-6. Plots of data set J in the reduced feature space.

data and resulted in the polytopes and rms proportion errors shown in Figure VIII-7. The raw data to reduced feature space transformation matrix for data set J is

$$TV = \begin{bmatrix} 0.606 & -0.021 & 0.127 \\ -0.290 & 0.589 & 0.471 \\ 0.485 & -0.067 & 0.342 \\ -0.094 & 0.588 & 0.053 \\ 0.185 & -0.069 & 0.674 \\ 0.521 & 0.545 & -0.434 \end{bmatrix} \quad (VIII - 11)$$

The end members plotted in Figure VIII-7 are given below.

Imbrie end members:

$$\begin{aligned} \vec{E}_{FY} &= [0.890 \quad 0.386 \quad 0.226] \\ \vec{E}_{SY} &= [0.378 \quad 0.296 \quad 0.873] \\ \vec{E}_{MY} &= [0.116 \quad 0.991 \quad 0.060] \end{aligned} \quad (VIII - 12)$$

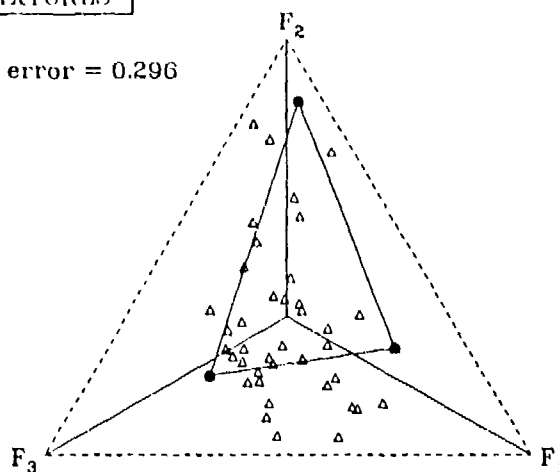
DENEG end members with $t_2 = -0.38$:

$$\begin{aligned} \vec{E}_{FY} &= [0.916 \quad 0.003 \quad 0.402] \\ \vec{E}_{SY} &= [0.429 \quad 0.235 \quad 0.873] \\ \vec{E}_{MY} &= [-0.025 \quad 0.996 \quad 0.083] \end{aligned} \quad (VIII - 13)$$

A DENEG polytope that encompasses the entire data cloud could not be achieved in this case. The DENEG iterative algorithm did not converge to a reasonable polytope when t_2 was large enough to include all data points. Imbrie end members, therefore, are the only valid comparison between these data and the corresponding data based on mixtures with uniform mean brightness. For varying means the unmixing accuracy is 0.296 compared to 0.275 (Figure VIII-2) for mixtures with preadjusted means. The difference between these two accuracy figures is small as expected from the observed similarity in raw data.

• IMBRIE EM
 ▲ MIXTURES

rms error = 0.296



• DENEG EM
 ▲ MIXTURES

rms error = 0.295

$t_2 = -0.38$

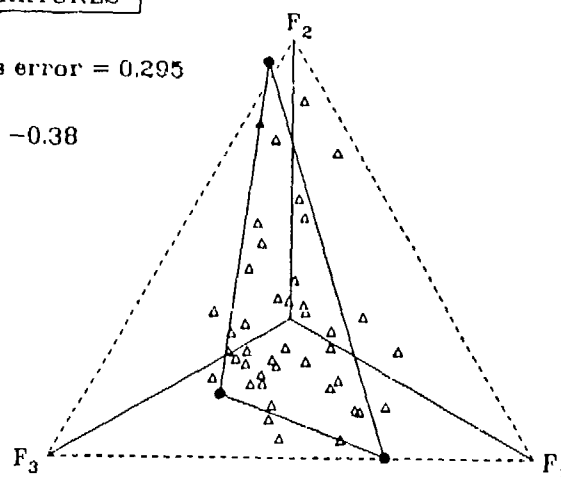


FIGURE VIII-7. Plots of data set J in the reduced feature space with various superimposed mixing polytopes.

2. Texture Plus MEAN

Figure VIII-8 is a plot of data set K. Again, Imbrie end members and DENEK end members are applied and the resulting polytopes are shown in Figure VIII-9. The transformation matrix from CABFAC for data set K is

$$TV = \begin{bmatrix} 0.901 & 0.046 & 0.090 \\ -0.033 & 0.698 & -0.403 \\ 0.421 & 0.083 & -0.257 \\ -0.096 & 0.548 & -0.163 \\ -0.002 & 0.452 & 0.858 \end{bmatrix}. \quad (VIII - 14)$$

End members from QMODEL for this case are as follows.

Imbrie end members:

$$\begin{aligned} \vec{E}_{FY} &= [0.992 \quad 0.109 \quad -0.050] \\ \vec{E}_{SY} &= [0.691 \quad 0.659 \quad 0.283] \\ \vec{E}_{MY} &= [-0.061 \quad 0.995 \quad -0.020] \end{aligned} \quad (VIII - 15)$$

DENEK end members with $t_2 = -0.3$:

$$\begin{aligned} \vec{E}_{FY} &= [0.997 \quad 0.064 \quad -0.053] \\ \vec{E}_{SY} &= [0.668 \quad 0.422 \quad 0.614] \\ \vec{E}_{MY} &= [-0.138 \quad 0.988 \quad -0.076] \end{aligned} \quad (VIII - 16)$$

Again, attempts at DENEK end members that encompassed all samples did not end in convergence, so only Imbrie unmixing is compared with previous results. Here when texture and MEAN are considered jointly and background intensity is allowed to vary in the mosaic, the Imbrie unmixing accuracy is 0.164, which is once again very close to the 0.157 value (Figure VIII-4) that resulted when means were preadjusted. Our conclusion is that there is no significant loss in accuracy of the calculated proportions when the mean intensity is allowed to vary. The objective of this section is therefore satisfied.

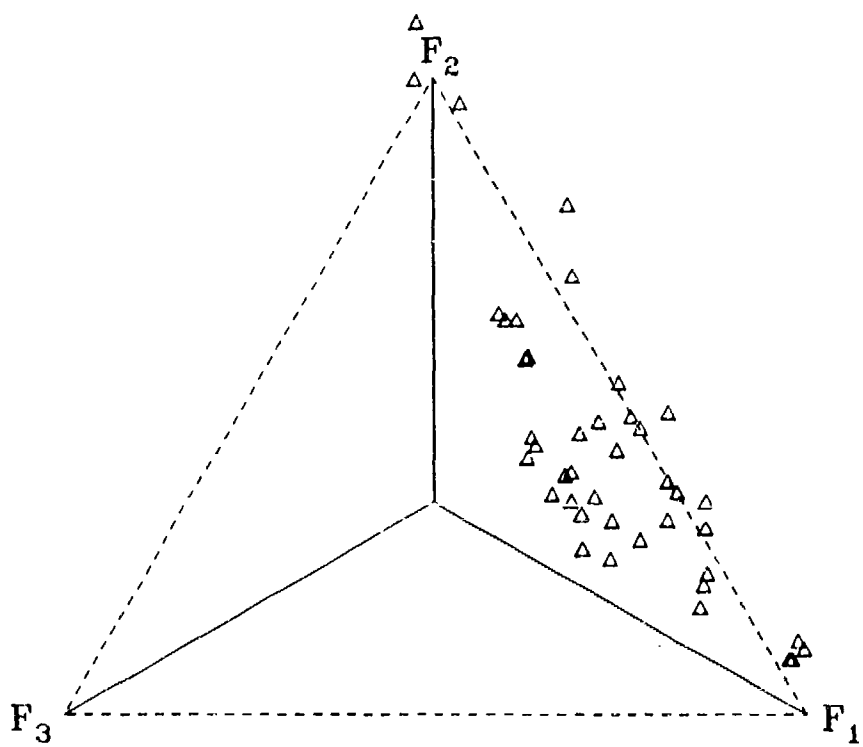
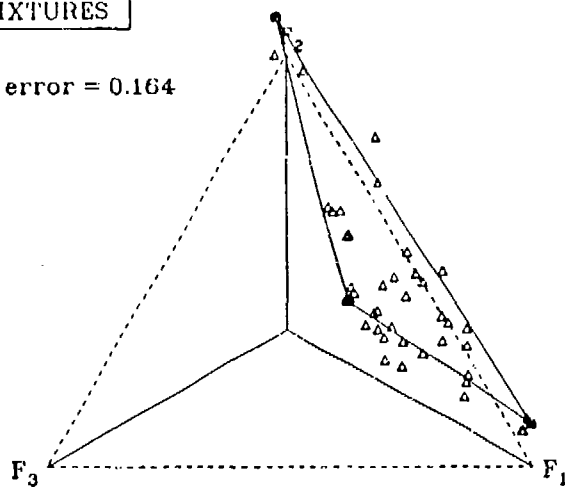


FIGURE VIII-8. Plots of data set K in the reduced feature space.

• IMBRIE EM
 △ MIXTURES

rms error = 0.164



• DENEG EM
 △ MIXTURES

rms error = 0.249

$t_2 = -0.30$

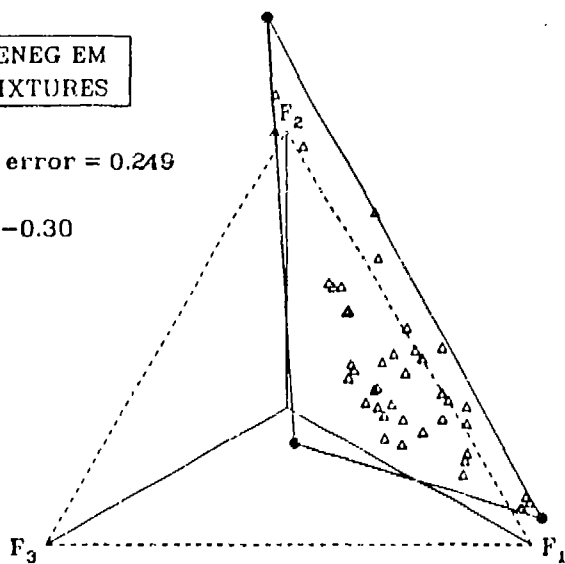


FIGURE VIII-9. Plots of data set K in the reduced feature space with various superimposed mixing polytopes.

3. Unmixing with All Variables of Data Set I

Although accuracy values do not indicate any problem with data set K, the shape of the data cloud in that case is troublesome. Note the Imbrie polytope in Figure VIII-9. The triangle is relatively collapsed so that the second-year end member is approaching the midpoint of the first-year/multiyear mixing line. This same problem (but more severe) was encountered in Section VI (see Figure VI-3). In the previous encounter we found that as long as either the LM4 and LM5 pair or MEAN was included in the data set, the polytope would expand to a more acceptable shape. In Figure VIII-7, when LM4 and LM5 are present, a nearly equilateral polytope results when Imbrie end members are connected. Figure VIII-9 does have MEAN included which has previously been sufficient for a solid three-end-member polytope, but it apparently is not sufficient in this case. It seems appropriate in view of this problem to include LM4, LM5, and MEAN in the same data set to see if an improved polytope can be obtained. This combination of variables was analyzed and discarded in Section VI because of the noisy nature of LM4 and LM5. However, since we have gone to larger images, the noise levels in the data are reduced, so another look at using all variables is in order.

Data set L is a SIFTed version of data set I with all 7 variables included. These data are listed in Table N-1 in Appendix N. Figure VIII-10 is a plot of the data. The data cloud in this figure tends to be more two-dimensional, and Figure VIII-11 shows an expanded polytope for the Imbrie end-member case. However, the accuracy is worse (0.231) for this expanded polytope than for the collapsed case in Figure VIII-9 (0.164). Apparently, the degradation from noise in LM4 and LM5 exceeds the benefits resulting from an improved polytope shape.

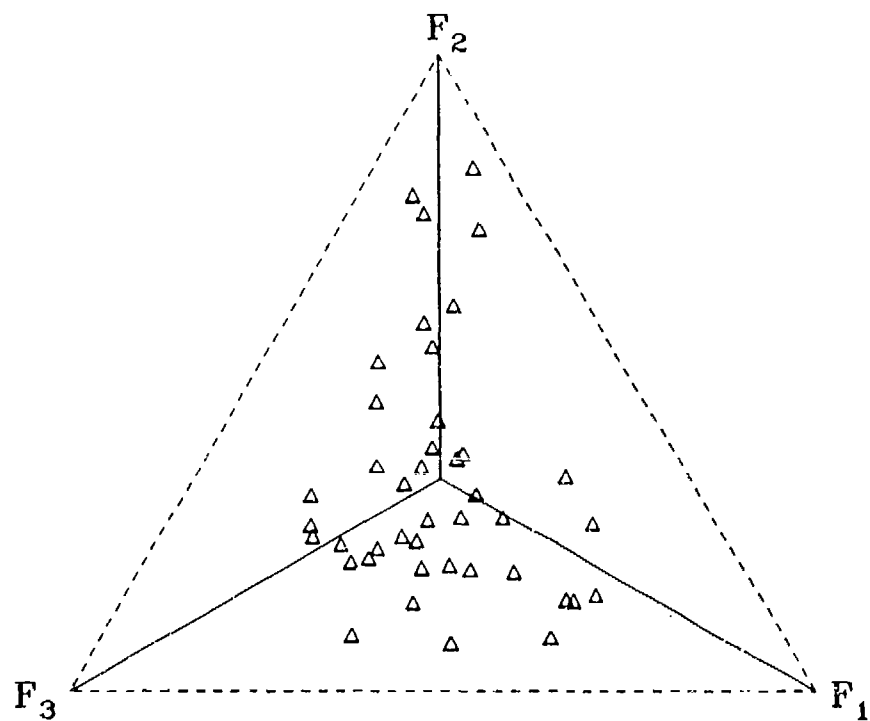
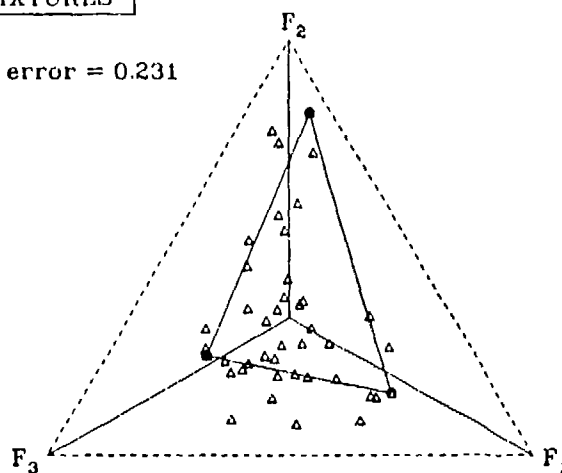


FIGURE VIII-10. Plots of data set L in the reduced feature space.

• IMBRIE EM
△ MIXTURES

rms error = 0.231



• DENEG EM
△ MIXTURES

rms error = 0.224

$t_2 = -0.40$

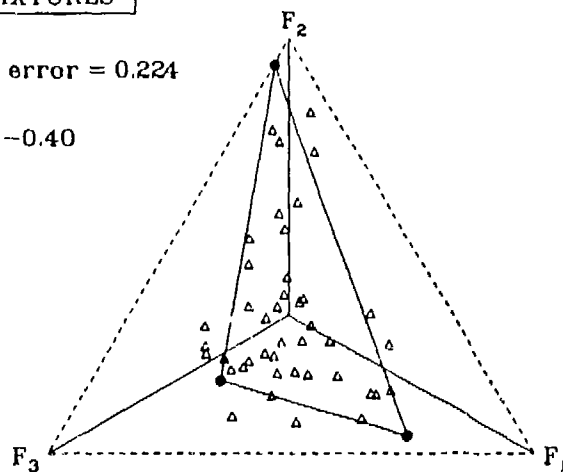


FIGURE VIII-11. Plots of data set L in the reduced feature space with various superimposed mixing polytopes.

Many texture investigators (*e.g.*, Laws, 1980) stop at this point, having demonstrated their algorithms on synthetic mixtures of image texture. Indeed, the primary objective of this study has been realized. The global approach to image texture has been developed and has been successfully applied to test imagery. However, analysis of complex natural texture composites, such as those encountered in remote sensing, is not an easy problem. Many techniques demonstrated on synthetic textures cannot be readily applied to practical applications. The complex issues of practical image texture analysis cannot be addressed within the scope of this work. Even given the time, the present KRMS data set would not be adequate to completely illustrate the global approach to analysis of natural mixtures. However, we do want to at least introduce the natural mixture problem, do some very limited work with this type of mixture, and point the way toward further work required to progress in this area. Section IX covers this topic.

IX. NATURAL MIXTURES

A. SIMPLE MIXTURES

When 512×312 pixel full frame KRMS images are considered, the types of mixtures available in this data set will not support a complete global texture analysis. There are no images in the data set that consist of only a single ice type. Therefore, we have no way to locate Imbrie end members for unmixing via extreme samples. Likewise, there are no images totally void of first-year ice. DENEK end members, therefore, cannot be defined using edge-extreme methods. However, four images containing simple mixtures of first-year and multiyear ice are available. Some insight into natural mixtures is available by analysis of these four images shown in Figures IX-1 through IX-4.

1. Building the Polytope from Large Synthetic Mixtures

One possibility to explore is to fill in the data set with some large synthetic mixtures. These mixtures could be used to fill in the holes in the polytope that cannot be defined from a limited supply of natural mixtures. These images should be 512×312 pixels like the natural mixtures to reduce noise levels, since the polytope will be based on so few points. Six such large synthetic mixtures, labeled BX5 through BX10, were constructed from the synthetic mixtures of Section VIII. Full 512×312 mosaics were not possible with the data at hand, but large mosaics ranging from 256×256 up to 512×256 pixels were constructed for each pure ice type and for 50-50 mixtures of FY/SY, FY/MY, and SY/MY. The six large synthetic mixtures, therefore, defined all three vertices and all three edges of the three-end-member polytope. The hope was to superimpose the polytope thus defined over the natural mixture images and calculate ice-type proportions on that basis. Figure IX-5 is the large multiyear mosaic, which

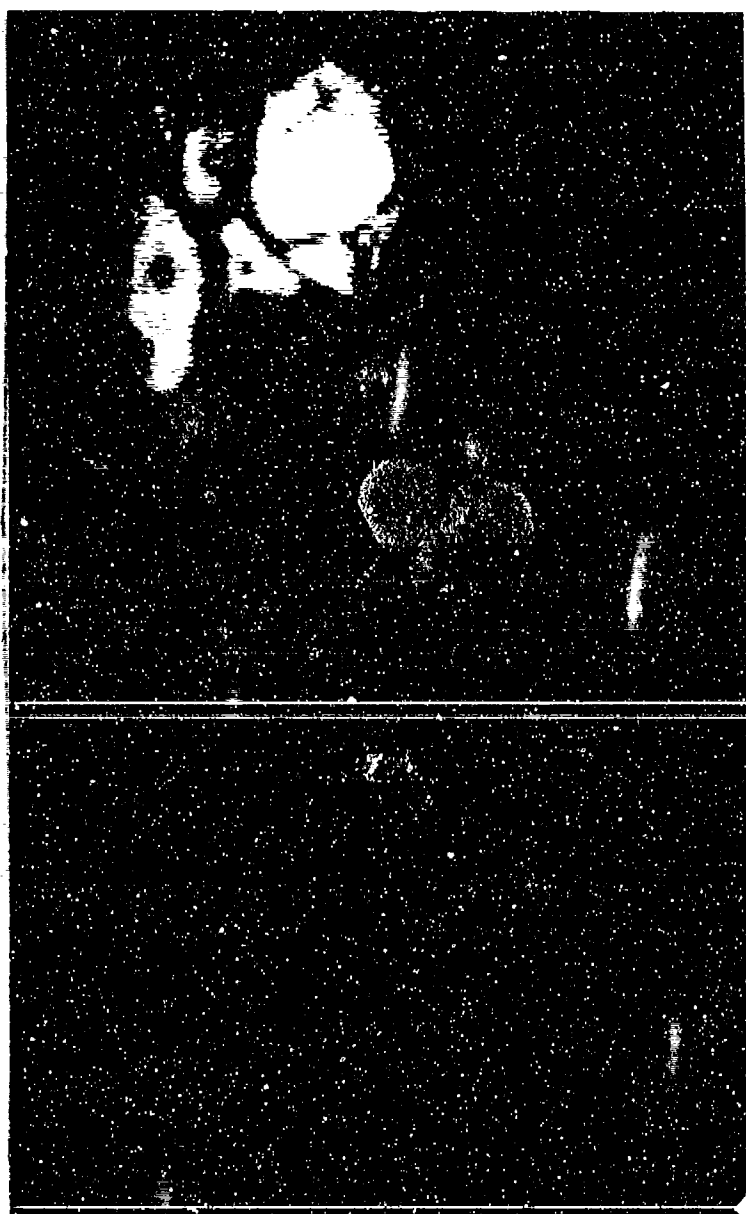


FIGURE IX-1. KRMS image BX1 containing a natural mixture of FY (90%) and MY (10%) ice. Image size is 312 samples by 512 lines.

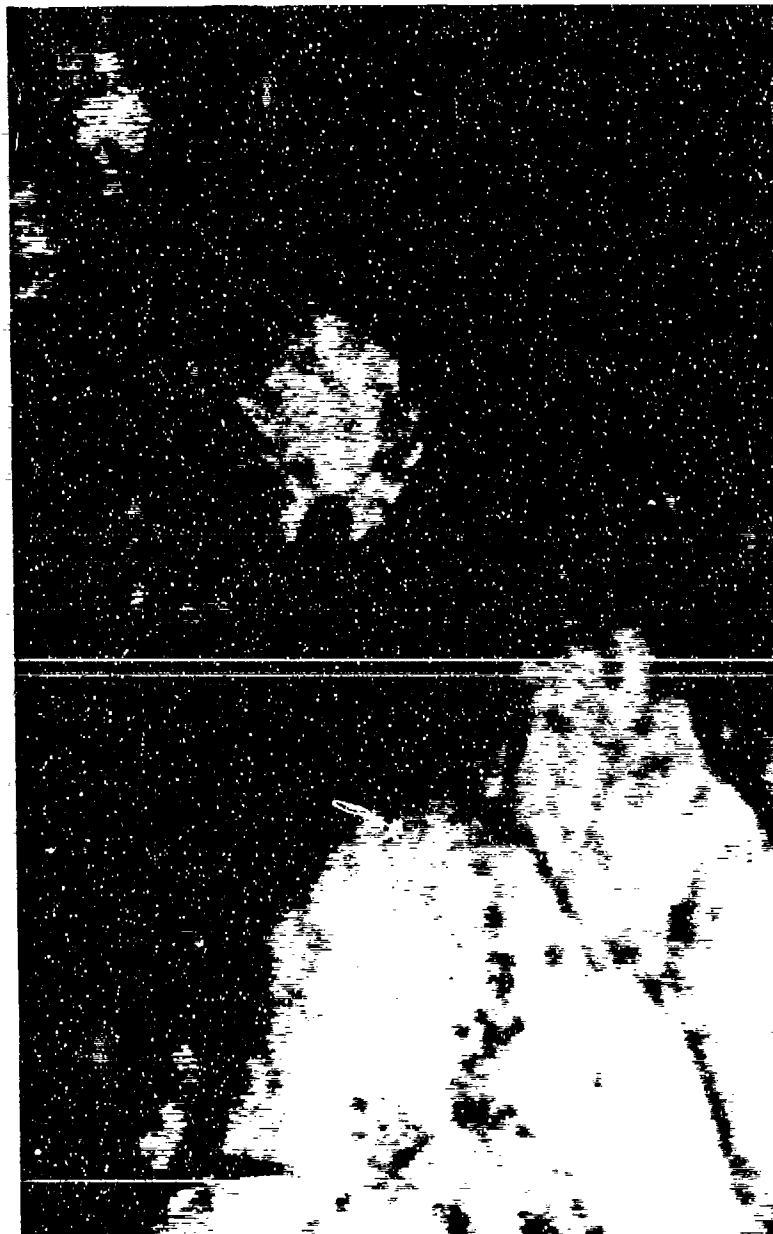


FIGURE IX-2. KRMS image BX2 containing a natural mixture of FY (40%) and MY (60%) ice. Image size is 312 samples by 495 lines.

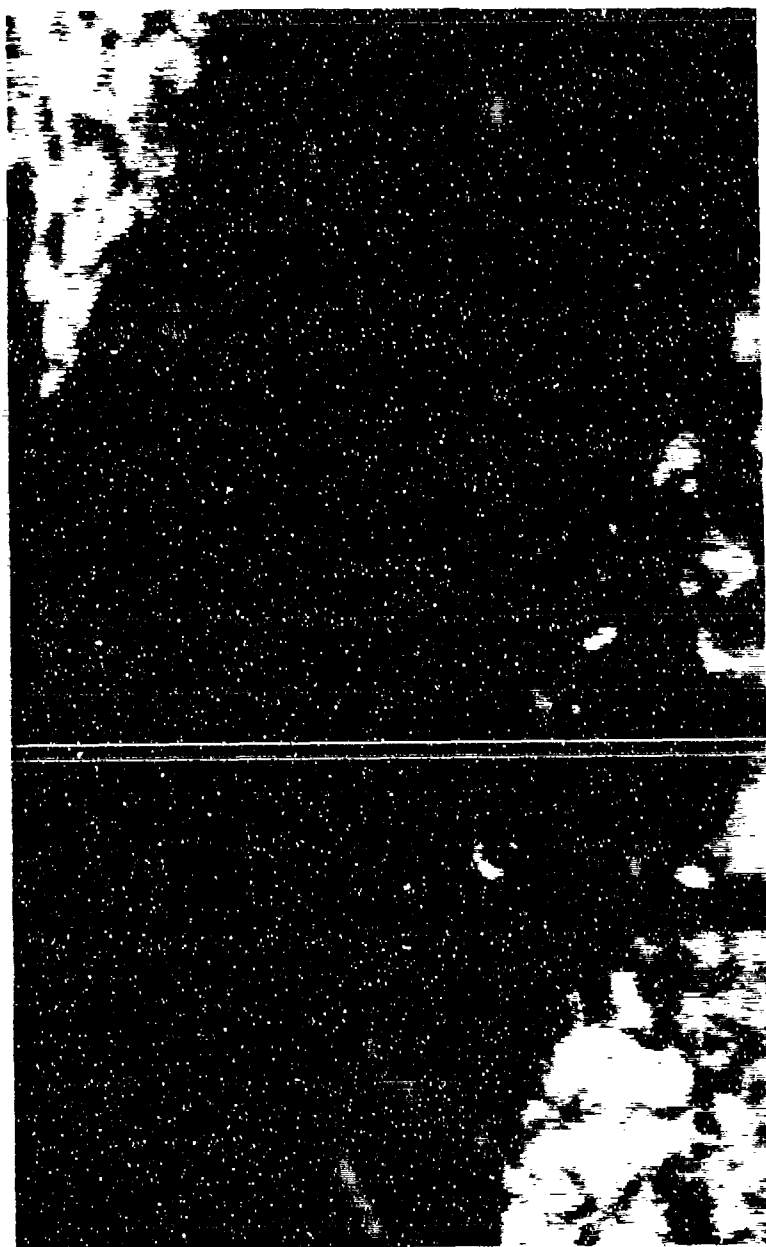


FIGURE IX-3. KRMS image BX3 containing a natural mixture of FY (69%) and MY (31%) ice. Image size is 312 samples by 512 lines.

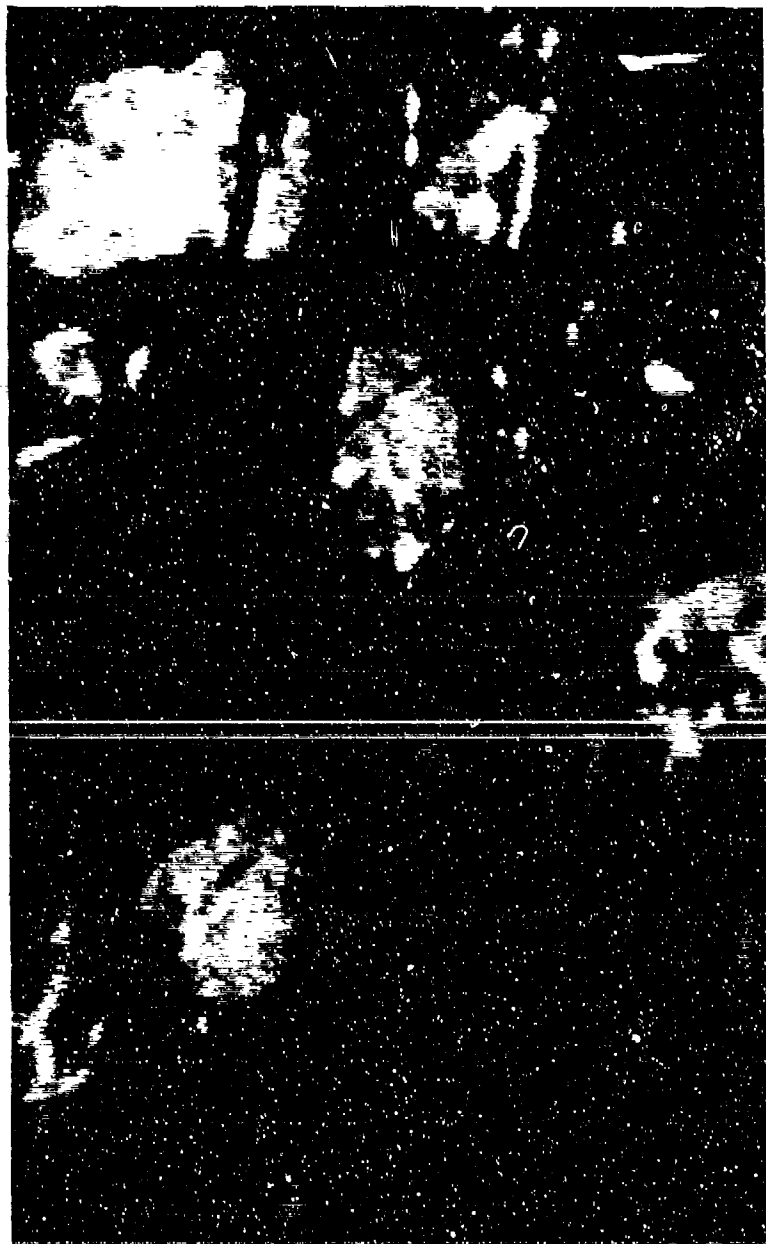


FIGURE IX-4. KRMS image BX4 containing a natural mixture of FY (60%) and MY (40%) ice. Image size is 312 samples by 512 lines.



FIGURE IX-5. Synthetic KRMS mixture image consisting entirely of MY ice. Image size is 256 samples by 384 lines.

is shown as a representative example of images BX5 through BX10. Table O-1 in Appendix O gives the calculated values of texture measures for these large images, which have named data set M.

It is most informative here to first consider texture variables HOM3, HOM15, LM2, and LM3, plus MEAN. These five variables have been extracted from Table O-1, SIFTed, and placed in Table P-1 (Appendix P) as data set N. CABFAC was used as before to transform data set N into a three-dimensional feature space. Figure IX-6 is a plot of data set N in this reduced feature space. In the figure the polytope constructed from the large synthetic mixtures is shown. The first observation is that the natural mixtures (triangular points in Figure IX-6) do not fall on the FY/MY mixing line as expected. They tend to fall in a line (see visual linear fit represented by the dotted line), but the mixing line for natural mixtures is approximately parallel to, but offset from, the FY/MY edge of the polytope.

An offset is not the only difference illustrated in Figure IX-6. The four natural-mixture points span multiyear concentrations of 10% to 60% or a range of 50%. The FY/MY edge of the polytope spans a range of 100% in multiyear concentration, yet its distance on the plot is approximately the same as for a 50% range of natural mixtures. Therefore, a scale, as well as an offset difference, seems to exist between texture measures calculated from synthetic and natural mixtures, so the idea of unmixing natural mixtures using a synthetic mixture polytope must be abandoned.

It is interesting to note how precisely positioned the six synthetic points are with respect to the polytope. With image sizes as large as these, the excessive noise we saw initially when working with 64×64 pixel images has been virtually

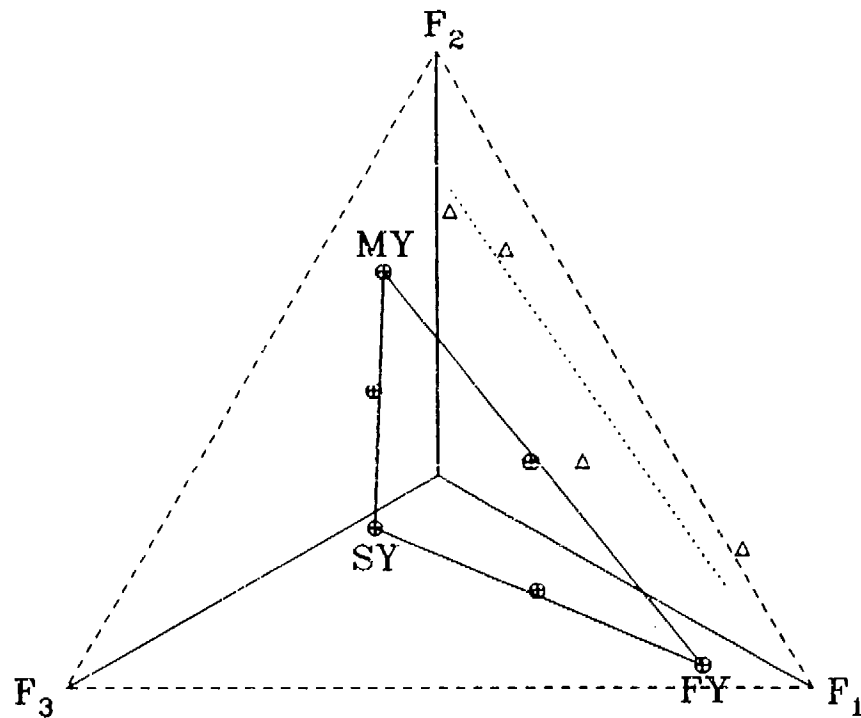


FIGURE IX-6. Plots of data set N in the reduced feature space. Triangles represent natural FY/MY mixture images. Crossed circles represent large synthetic mixtures. Polytope is constructed by connecting large synthetic mixtures of a single ice type.

eliminated. The polytope has been formed by connecting the pure ice type vertices exactly. A comparison of the edge points with their expected positions midway between vertices will give an indication of the unmixing accuracy that can be achieved with these larger images, which, on the average, were equivalent to 300×300 pixels. In Figure IX-6 the calculated proportions agreed with the known proportions of the edge points to an rms error of 0.039.

2. Two-End-Member Analysis of the Natural Mixtures

Since the natural mixtures BX1 through BX4 do fall in a line, one might get some estimate of possible unmixing accuracy by attempting unmixing calculations along that line. We have no end members for the dotted mixing line shown in Figure IX-6. However, one can determine if the relative positions along that line are related in a consistent manner to first-year/multiyear ice proportions. To determine whether consistency is present, each sample's position was projected onto the dotted mixing line, and that position was measured relative to some reference point on the line. For convenience, the lower sample in Figure IX-6 was chosen as the reference point. Measured distances for the four points were 0.0, 0.87, 2.00, and 2.32 inches, which corresponded to multiyear proportions of 0.10, 0.31, 0.40, and 0.60. A linear least-squares curve fit was applied to these four data points, resulting in the equation.

$$MY \text{ proportion} = (\text{distance} \times 0.1857) + 0.1116. \quad (IX - 1)$$

This equation, being a best fit relationship of mixing line distance to multiyear proportion, can be used to estimate proportions based on distance. The calculated proportions from Eq. (IX-1) for the four samples in question were 0.112, 0.273, 0.483, and 0.542 leading to an rms proportion error of 0.054 for unmixing

of these four samples. It would be wrong to place too much emphasis on accuracy values based on four points. However, it is interesting to note that this value is close to the accuracy value of 0.05, which would be predicted for these large images based on the 128×128 synthetic mixtures in Section VIII, plus sample size considerations. We therefore have some evidence that these simple natural mixtures are not significantly more difficult to unmix than were the synthetic mixtures.

The same analysis of data set M was attempted using texture measures only. The six texture measures were SIFTed to form data set O which is tabulated in Appendix Q. Figure IX-7 is a plot of that data. Again, as in Figure IX-6, a polytope based on large synthetic mixtures BX5 through BX10 is plotted, as well as a visual fit of a mixing line through the natural samples. In this texture-only case, the natural mixing line and the FY/MY side of the polytope are very close. The offset observed in Figure IX-6 is not present here. However, the scaling discrepancy remains. The polytope spans a 100% range of concentrations in approximately the same distance as is covered by a 50% concentration range in natural mixtures.

Unmixing of the natural mixture samples was attempted again using procedures previously described. In Figure IX-7 the measured distances were 0.0, 0.79, 1.86, and 2.29 inches. These measurements led to the following least-squares predictor of multiyear proportions.

$$MY \text{ proportion} = (\text{distance} \times 0.1926) + 0.1146. \quad (IX - 2)$$

Resulting unmixed proportions of multiyear ice were (via Eq. (IV-2)) 0.115, 0.267, 0.473, and 0.556, leading to an rms proportion error of 0.048. This unmixing accuracy is well below the value of 0.095, which would be predicted for these

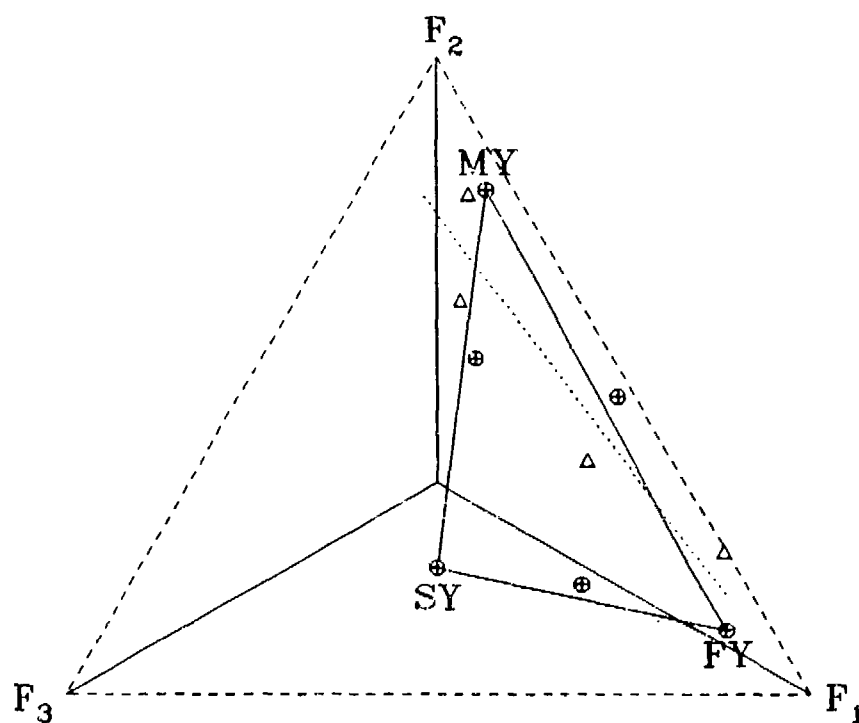


FIGURE IX-7. Plots of data set O in the reduced feature space. Triangles represent natural FY/MY mixture images. Crossed circles represent large synthetic mixtures. Polytope is constructed by connecting large synthetic mixtures of a single ice type.

images based on texture-only results in Section VIII. Texture-only accuracy here is as good as the texture plus MEAN case. This equality of the two cases is inconsistent with results in previous chapters where texture only was worse than texture plus MEAN. However, little significance rests with numbers based on four samples. This limited study of simple natural textures indicates that simple natural textures can possibly be as easily and accurately unmixed as synthetic mixtures.

We can look again at how well the polytope sides in Figure IX-7 match the 50-50 synthetic mixture samples that should fall at the midpoints of the sides. For texture only in Figure IX-7, the calculated proportions agreed with the known proportions of the edge points to an rms error of 0.074.

3. Unmixing Accuracy as a Function of Image Size

In Section VI the assumption was made that unmixing accuracy for global texture analysis would vary as 1 over the square-root of the number of pixels in the image, or in other words, as 1 over the dimension of a single side of a square image. That assumption was put forward based on the facts that all variables were normal and that all unmixing procedures are linear. At this point we have data available to confirm this assumption. We have error values for 64×64 images from Section VI, 128×128 images from Sections VII and VIII, 512×312 natural images, and 256×384 synthetic images in this section. If we take the best texture plus MEAN error values obtained for each image size (whether Imbrie, DENEG, or class mean end members) and plot these accuracy values as a function of image size, then the plot shown in Figure IX-8 is obtained. For comparison, a 1 over sample size curve is shown as a dashed line. A perfect match with the dashed curve would not be expected, since some of the square

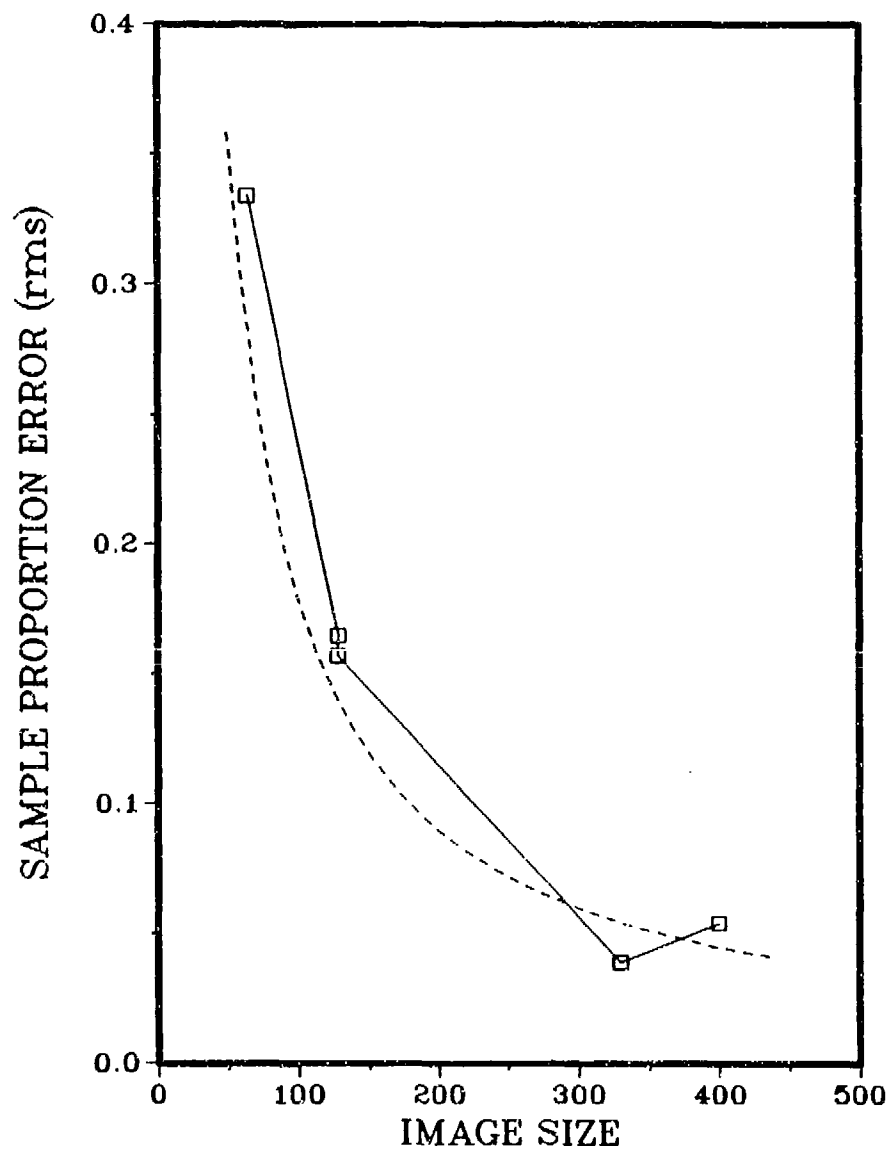


FIGURE IX-8. Calculated errors in unmixing proportions as a function of image size, S . The dashed curve is a S^{-1} reference line.

points are for pure samples while others are for mixtures, and some mixtures are synthetic while others are natural. Also the relative representations of each texture type is different for each data set. In spite of these differences, the points fall close enough to a S^{-1} behavior, where S is the length of the image side, that this assumed relationship to image size can be considered experimentally confirmed.

B. COMPLEX MIXTURES

Texture types in natural images will not always occupy clearly defined areas with simple boundaries, as was the case in Figures IX-1 through IX-4. Figure IX-9 is an example of such a mixture. Note the two boxes drawn on the figure. Both boxes include multiyear ice. However, the upper box is positioned on a large solid multiyear floe, while the lower box contains numerous small pieces of multiyear ice. Clearly the fragmented nature of the ice in the lower box imparts an entirely different textural appearance than is present in the upper box, even though both are filled with multiyear ice. When the ice mixtures become complex through fracturing or through formation of rubble fields consisting of small chunks of assorted ice types, ice classification based on image texture would be expected to encounter problems.

These problems are not believed to be insurmountable. One solution is suggested and illustrated here. The proposal is to preprocess the image to identify areas of complex textures. These areas can then be replaced by imbedding "blotches" of uniform gray to replace the complex areas. When the unmixing is performed, these uniform areas will fall out of the analysis as an additional end member. That is, rather than three end members representing first-year, second-year, and multiyear ice, CABFAC would indicate a four-dimensional fea-



FIGURE IX-9. Complex KRMS natural mixture of FY and MY ice.
Boxes denote a solid MY flow and MY rubble.

ture space where three of the end members represent ice types and the fourth end member, which could be called "unclassifiable" would represent the previously masked areas of complex structure.

An alternate approach would be to just leave the complex areas as they are and let the unmixing procedure assign end members to these directly. The author does not believe that this approach would be satisfactory. The wide variety of types of complex texture would require numerous end members which would lead to a difficult interpretation of the results. It seems preferable to preprocess and join all complex, unclassifiable areas into a single texture (uniform gray) and then proceed with unmixing.

To show that such a preprocessing step might be feasible, the following steps were applied to Figure IX-9. First, a cluster shade edge detector (Holyer and Peckinpaugh, 1989) was applied to the image with an edge threshold that detected fractures and rubble but did not have the sensitivity to detect texture within floes. (This new edge detector was an outgrowth of the GLC matrix-based texture analyses described in Section III). Edges thus detected were dilated for several iterations to make the edge lines wider. The dilation process nearly filled in complex areas with high edge density such as a rubble field. All image pixels under the dilated edge mask are set to a uniform mid-range grey value (in this case, 128). Figure IX-10 shows the resulting image. Rubble areas have been eliminated and solid floes remain. An unmixing analysis of this image should give accurate proportions for the first-year ice and the remaining first-year floes, plus one more proportion for the masked areas of unmanageable complexity. What has been shown here as an example is a simple preprocessing step. Certainly better methods for masking complex ice areas could be developed. Perhaps something as simple as variance in a local neighborhood could be used to delineate complex

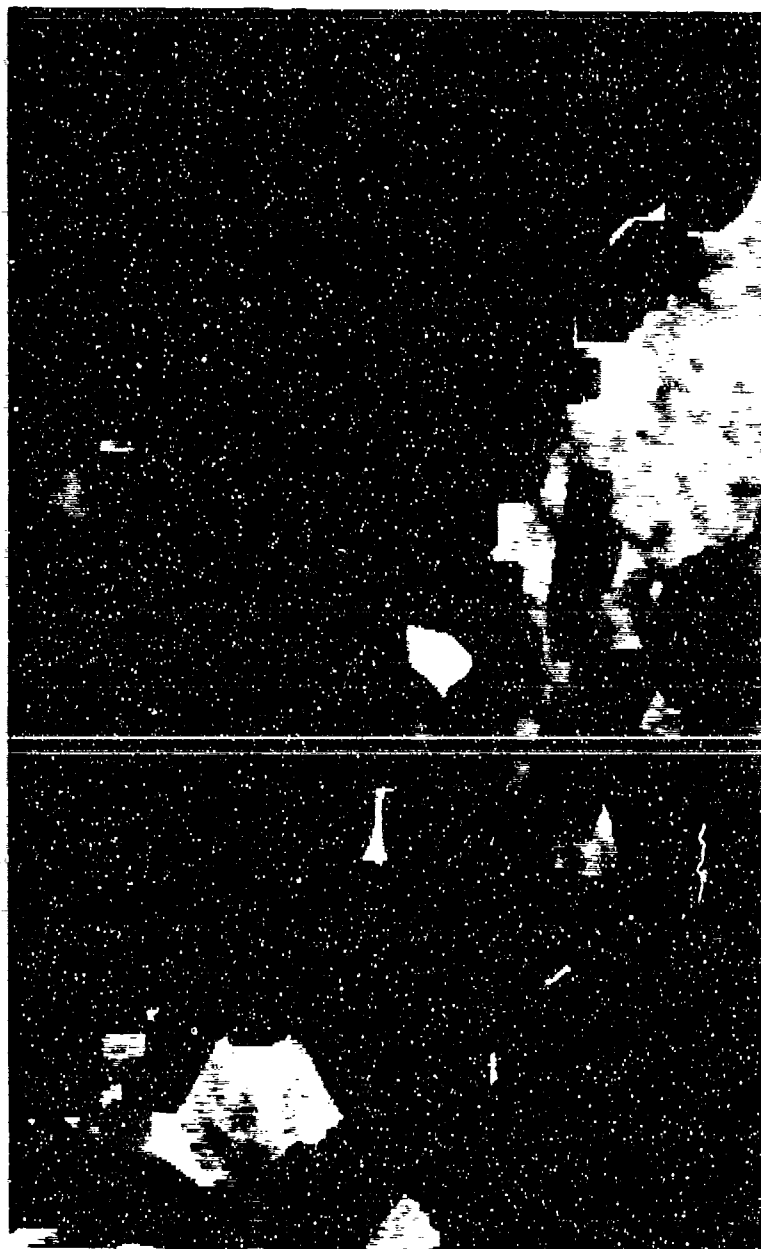


FIGURE IX-10. Natural mixture image of Figure IX-9 with complex rubble areas covered with a uniform gray level.

textural areas.

X. CONCLUSIONS AND RECOMMENDATIONS

Conclusions resulting from this study will be divided into two categories. The first deals with the global texture analysis paradigm, which was the major thrust of this work. These conclusions cover matters of broad applicability to the field of image processing. The second category deals with information learned about sea ice as a result of using imagery of that type as test data.

A. THE GLOBAL TEXTURE ANALYSIS PARADIGM

- The global approach to texture analysis has been developed and applied to a test data set with sufficiently promising results that it can be concluded that the approach may have merit for many image texture problems.
- Both mandatory and desirable characteristics of texture measures were enumerated. Of 25 candidate texture measures evaluated, none met all of the desirable and mandatory characteristics, and only two – local homogeneity and Laws' energy masks – possessed the mandatory characteristics, which included correlation with ice type and linear behavior under formation of mixtures. Here, of course, the key item is the linear mixing property. Investigation of additional texture measures that exhibit this characteristic would support further exploitation of the global texture analysis approach.
- The SIFT procedure was developed and applied here to deal with the nonconstant-sum nature of image texture variables. Since this work was performed, a variable-sum version of CABFAC and QMODEL is reportedly available (Full, 1988). The analysis described here should be repeated using these new programs. If the SIFT used here resulted in significant loss of ice-type information, the repeated analysis with the new code would give better results, indicating even more promise for the global techniques.

- The global approach should be applied to other texture data sets to see if it does indeed have broad applicability.
- Further study should be initiated to deal with the problem of complexity occurring in natural images.
- Failure to solve the complexity problem does not eliminate the global method from practical application. There may be certain applications, industrial inspection and control, for example, where the scenes are always simple mixtures and this technique is applicable without presimplification.
- Global unmixing of image texture is inexorably linked to a trade-off. Proportion accuracy can be exchanged for positional uncertainty. If the global area is large, proportions can be extracted quite accurately, but one does not know where within the global area the particular texture in question is located. Conversely, as the global area is made smaller to pinpoint the location of a feature, the proportion filled by that texture feature can be measured with less certainty. This situation suggests that a pyramid approach could be advantageously applied. Large areas could be analyzed for maximum accuracy followed by analysis of progressively smaller areas in selected parts of the image where it is judged to be advantageous based on criteria that would be unique to a given application.

B. SEA ICE SCIENCE

- Proportions of first-year, second-year, and multiyear ice types in samples extracted from KRMS images can be calculated with useful accuracy using linear unmixing of globally determined texture measures.
- Eppler *et al.* (1984) studied classification of ice types in the same KRMS

data set. Their classification work involved thresholding on image brightness. They concluded that brightness temperature thresholding alone was not adequate for classification of ice types. The recommendation in that report was that image texture be included, and provided the stimulus for the present study. We demonstrated here that texture can lead to calculation of accurate (accuracy depending on image size) ice-type proportions. Furthermore, the accuracy has been shown to improve by an approximate factor of two when mean brightness levels are added to texture data. Unfortunately, Eppler *et al.* (1984) did not give any quantitative estimates of image composition accuracies resulting from the thresholding method, so that direct comparison could be made with the present results.

- The proportion errors as a function of image size that have been derived here indicate what order of spatial scales are required to adequately characterize ice types. The 312×512 pixel images that resulted in proportion errors of approximately 0.05 represent 6.4 km^2 of ice surface.
- The most critical hurdle to practical application of the global technique to ice analysis is believed to be the development of a preprocessing step that will identify and simplify areas of complex textural structure.
- Assuming the above can be accomplished, this texture analysis scheme should be applied to another KRMS data set, where adequate representation of the various ice types will permit complete unmixing using only natural images.
- The real future of ice-texture analysis lies with SAR rather than with passive microwave sensors. The KRMS is a research tool, but operationally, the SAR is the sensor of the future. The ERS-1 satellite will have a SAR in space in 1991 and for the foreseeable future thereafter. NASA is establishing

a receiving facility in Fairbanks, Alaska, to receive that data over a large portion of the Arctic. The Navy Polar Oceanography Center in Suitland, Maryland, is planning to use this SAR data to generate its operational ice products for the U.S. Navy. Clearly, if the texture analysis techniques developed here could be extended to SAR imagery, then the potential payoff would be large in terms of practical application.

XI. REFERENCES

- BAJCSY, R. and L. LIEBERMAN. Computer description of real outdoor scenes. *Proceedings, Second International Conference on Pattern Recognition*, Copenhagen, Denmark, 1974, 174-179.
- BAJCSY, R. and L. LIEBERMAN. Texture gradient as a depth cue. *Comput. Graphics and Image Processing* 5, 1, 1976, 52-67.
- BALLARD, D. H. and C. M. BROWN. *Computer Vision*. Englewood Cliffs, N.J., Prentice-Hall, 1982.
- BEZDEK, J. Numerical taxonomy with fuzzy sets. *Jour. Math. Bio.* 1, 1974, 57-71.
- BEZDEK, J. *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York, Plenum Press, 1981.
- BRACEWELL, R. *The Fourier Transform and Its Applications*. New York, McGraw-Hill, 1965.
- BRODATZ, P. *Textures: A Photographic Album for Artists and Designers*. New York, Dover, 1966.
- BURY, K. V. *Statistical Models in Applied Science*. New York, John Wiley and Sons, 1975.
- CONNERS, R. W. Towards a set of statistical features which measure visually perceivable qualities of texture. *Proceedings of IEEE Conference on Pattern Recognition and Image Processing*, Chicago, Illinois, August 1979, 382-390.

- CONNERS, R. W., M. M. TRIVEDI, and C. A. HARLOW. Segmentation of high-resolution urban scene using texture operators. *Comput. Graphics and Image Processing* 25, 1984, 273-310.
- CULLEN, C. G. *Matrices and Linear Transformations*. Reading, Mass., Addison-Wesley, 1966.
- DAVIS, L. S., S. A. JOHNS, and J. K. AGGARWAL. Texture analysis using generalized co-occurrence matrices. *IEEE Trans. PAMI-1*, 3, 1979, 251-259.
- DIXON, W. J. and F. J. MASSEY. *Introduction to Statistical Analysis*. New York, McGraw-Hill, 1983.
- DUDA, R. O. AND P. E. HART. *Pattern Recognition and Scene Analysis*. New York, Wiley, 1973.
- DUNN, W. J. The effect of inbreeding on the bones of the fowl. *Storrs Agricultural Experiment Station Bulletin* 52, 1928, 1-112.
- EHRLICH, R. and W. E. FULL. Sorting out geology - unmixing mixtures. *Use and Abuse of Statistical Methods in the Earth Sciences*. W. B. Size, Ed., New York, Oxford Univ. Press, 1987, 33-46.
- EPPLER, D. T. Personal communication. 1987.
- EPPLER, D. T. and D. FARMER. Personal communication. 1983.
- EPPLER, D. T., L. D. FARMER, A. W. LOHANICK, and M. HOOVER. *Digital Processing of Passive K_a -band Microwave Images for Sea-ice Classification*. Naval Ocean Research and Development Activity, Stennis Space Center, Miss., NORDA Report 51, 1984.

- FILLY, M. and D. A. ROTHROCK. Extracting sea ice data from satellite SAR imagery. *IEEE Trans. GE-24*, 1, 1986, 849-854.
- FINKBEINER, D. T. *Introduction to Matrices and Linear Transformations*. W. H. Freeman and Co., 1966.
- FULL, W. E. Personal communication. 1988.
- FULL, W. E., R. EHRLICH, and J. E. KLOVAN. EXTENDED QMODEL - objective definition of external end members in the analysis of mixtures. *Jour. Math. Geol.* 13, 4, 1981, 331-344.
- FULL, W. E., R. EHRLICH, and J. C. BEZDEK. FUZZY QMODEL - a new approach for linear unmixing. *Jour. Math. Geol.* 14, 3, 1982, 259-270.
- GAGALOWICZ, A. Stochastic texture synthesis from *a priori* given second-order statistics. *Proceedings IEEE Computer Society, Conference on Pattern Recognition and Image Processing*, Chicago, Illinois, August 1979, 376-381.
- GAGALOWICZ, A. A new method for texture field synthesis: some applications to the study of human vision. *IEEE Trans. PAMI-3*, 5, 1981, 520-533.
- GALLOWAY, M. M. Texture analysis using gray level run lengths. *Comput. Graphics and Image Process.* 4, 1975, 172-179.
- HARALICK, R. M. Statistical and structural approaches to texture. *Proceedings IEEE* 67, 5, 1979, 786-804.
- HARALICK, R. M., K. SHANMUGAM, and I. DINSTEIN. Textural features for image classification. *IEEE Trans. SMC-3*, 6, November 1973, 610-621.
- HARLOW, C. A. and R. W. CONNERS. *A Study of Texture Analysis Algorithms*. Remote Sensing and Image Processing Laboratory, Louisiana

State University, Baton Rouge, Louisiana, Final Technical Report, Grant No. AFOSR-18-0112, 1983.

HARWOOD, D., S. MURALIDHARA, and L. S. DAVIS. Texture classification by local rank correlation. *Comput. Graphics and Image Processing* 32, 1985, 404-411.

HAYDEN, C. H., R. C. GONZALEZ, and A. PLOYSONGSANG. A temporal edge-based image segmentor. *Pattern Recognition* 20, 3, 1987, 281-290.

HOLMES, Q. A., D. R. NÜESCH, AND R. A. SHUCHMAN. Textural analysis and real-time classification of sea ice types using digital SAR data. *IEEE Trans. GE-22*, 2, 1984, 113-120.

HOLYER, R. J., and S. H. PECKINPAUGH. Edge detection applied to satellite imagery of the oceans. *IEEE Trans. GE-27*, 1, 1989, 46-56.

HSU, S. The Mahalanobis classifier with the generalized approach for automated analysis of imagery texture data. *Comput. Graphics and Image Processing* 9, 1979, 117-134.

IMBRIE, J. *Factor and Vector Analysis Programs for Analyzing Geologic Data*. Office of Naval Research, Geography Branch, Washington, D.C., Tech. Rept. 6, ONR Task No. 389-135, 1963.

IMBRIE, J. and T. H. VAN ANDEL. Vector analysis of heavy mineral data. *Bull. Geol. Soc. Amer.* 75, 1964, 1101-1110.

JOHNSON, R. A. and D. W. WICHERN. *Applied Multivariate Statistical Analysis*. Englewood Cliffs, N. J., Prentice-Hall, 1982.

JULESZ, B. Visual pattern discrimination. *IRE Trans. Info. Theory* IT-8, 2, 1962, 84-92.

- KAISER, H. F. The varimax criterion for analytic rotation in factor analysis. *Psychometrika* 23, 3, 1958, 187-200.
- KLOVAN, J. E. A generalization of extended Q-mode factor analysis to data matrices with variable row sums. *Math. Geol.* 13, 3, 1981, 217-224.
- KLOVAN, J. E. and J. IMBRIE. An algorithm and FORTRAN-IV program for large-scale Q-mode factor analysis and calculation of factor scores. *Math. Geol.* 3, 1, 1971, 61-77.
- KLOVAN, J. E. and A. T. MIESCH. EXTENDED CABFAC and QMODEL computer programs for Q-mode factor analysis of compositional data. *Comput. Geosci.* 1, 1976, 161-178.
- LANDERWEERD, G. H. and E. S. GELSEMA. The use of nuclear texture parameters in the automatic analysis of leukocytes. *Pattern Recognition* 10, 1978, 57-61.
- LAWS, K. I. Textured image segmentation. Ph.D. dissertation, Dept. of Engineering, Univ. Southern California, Los Angeles, California, 1980.
- LENDARIS, G. G. and G. L. STANLEY. Diffraction pattern samplings for automatic pattern recognition. *Proccings, IEEE* 58, 2, 1970, 198-216.
- LYDEN, J. D., B. A. BURNS, and A. L. MAFFETT. Characterization of sea ice types using synthetic aperture radar. *IEEE Trans. GE-22*, 5, 1984, 431-439.
- MACDUFFEE, C. C. *Vectors and Matrices*. The Mathematical Association of America, 1943.
- MALESON, J., C. BROWN, and J. FELDMAN. Understanding natural texture. University of Rochester, Rochester, New York, 1977.

- MANSON, V. and J. IMBRIE. FORTRAN program for factor and vector analysis of geologic data using an IBM 7090 or 7094/1401 computer system. Kansas Geol. Survey, *Sp. Dist. Publ. 13*, 1964.
- MIESCH, A. T. Q-mode factor analysis of compositional data. *Comput. Geosci.* 1, 3, 1976a, 147-159.
- MIESCH, A. T. Q-mode factor analysis of geochemical and petrologic data matrices with constant row sums. U.S. Geol. Survey Prof. Paper 574-G. 1976b.
- MIESCH, A. T. Interactive computer programs for petrologic modeling with extended Q-mode factor analysis. *Comput. Geosci.* 2, 4, 1976c, 439-492.
- MORRISON, D. F. *Multivariate Statistical Methods*. New York, McGraw-Hill, 1976.
- NASHBURG, R. E. and M. LINEBERRY. Noise effects for edge operators. *Proceedings, SPIE 292*, 1981, 277-287.
- PEET, F. G. and T. S. SAHOTA. Surface curvature as a measure of image texture. *IEEE Trans. PAMI-7*, 6, 1985, 734-738.
- PRATT, W. K., O. D. FAUGERAS, and A. GAGALOWICZ. Applications of stochastic texture field models to image processing. *Proceedings, IEEE 69*, 5, May 1981.
- PREWITT, J. M. S. Object enhancement and extraction. *Picture Processing and Psychopictorics*, 1970, 75-149.
- ROSENFELD, A. A note on automatic detection of texture gradients. *IEEE Trans. Comput.* 23, 10, 1975, 988-991.
- ROSENFELD, A. and M. THURSTON. Edge and curve detection for visual scene analysis. *IEEE Trans. Comput.* 20, 5, 1971, 562-569.

- STOGRYN, A. The brightness temperature of a vertically structured medium. *Radio Science* 5, 12, 1970, 1397-1406.
- STOGRYN, A. *Electromagnetic Properties of Random Media*. Aerojet Electric Systems Company, Azusa, California, Tech. Report TR-1, no. 6140, 1981.
- THRALL, R. M. and L. TORNHEIM. *Vector Spaces and Matrices*. New York, John Wiley and Sons, 1957.
- TRIENDL, E. E. Automatic terrain mapping by texture recognition. *Proceedings Eighth Internat. Symp. on Remote Sensing of Environment*, Environmental Research Institute of Michigan, Ann Arbor, Mich., 1972.
- TRIVEDI, M. M., C. A. HARLOW. Identification of unique objects in high resolution aerial images. *Optical Engineering* 24, 3, 1985, 502-506.
- TRIVEDI, M. M., C. A. HARLOW, R. W. CONNERS, and S. GOH. Object detection based on gray level cooccurrence. *Comput. Graphics and Image Processing* 28, 1984, 199-219.
- WERMAN, M. and S. PELEG. Min-max operators in texture analysis. *IEEE Trans. PAMI-7*, 6, 1985, 730-733.
- WESZKA, J. S., C. R. DYER, and A. ROSENFELD. A comparative study of texture measures for terrain classification. *IEEE Trans. SMC-6*, 4, April 1976, 269-285.
- WHITE, P. A. *Linear Algebra*. Dickenson Publishing Co., 1966.
- WILKINSON, J. H. Housholder's method for solution of the algebraic eigenproblem. *Computer Journal* 3, April 1960, 23-27.
- WMO. *WMO Sea-Ice Nomenclatures*. Secretariat of the World Meteorological Organization, Geneva, Switzerland, 1970.

APPENDIX A

THE UNIVARIATE MAXIMUM BAYESIAN CLASSIFIER

The univariate normal distribution, with mean μ and variance σ^2 , has the probability density function, $P(x)$,

$$P(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2}}, \quad (A-1)$$

where $z = \frac{(x-\mu)}{\sigma}$. If we consider l data sets each representing one of l possible normally distributed classes, a probability density function, $P_l(x)$, could be calculated for each of the l classes,

$$P_l(x) = \frac{1}{\sqrt{2\pi\sigma_l^2}} e^{-\frac{z_l^2}{2}}, \quad (A-2)$$

where μ_l and σ_l are the mean and standard deviation of the l^{th} class, respectively, and $z_l = \frac{(x-\mu_l)}{\sigma_l}$. The maximum Bayesian classification rule is to assign an unknown sample x_n to class l if

$$P_l(x_n) > P_m(x_n), \quad \text{for all } l \neq m. \quad (A-3)$$

The threshold on x separating classes l and m is given by the equation

$$P_l(x) - P_m(x) = 0. \quad (A-4)$$

The x value at which Eq. (A-4) is true can be considered the threshold, $T_{l,m}$, on x , above which sample x_n would be classified as belonging to class m and below which x_n would be assigned to class l . An expression for $T_{l,m}$ can be obtained by substituting expressions for $P_l(x)$ and $P_m(x)$ from Eq. (A-2) into Eq. (A-4). This procedure results in a quadratic form for T_{lm} .

$$aT_{lm}^2 + bT_{lm} + c = 0, \quad (A-5)$$

where

$$\begin{aligned} a &= \frac{1}{\sigma_m^2} - \frac{1}{\sigma_l^2} \\ b &= 2 \left(\frac{\mu_l}{\sigma_l^2} - \frac{\mu_m}{\sigma_m^2} \right) \\ c &= \frac{\mu_m^2}{\sigma_m^2} - \frac{\mu_l^2}{\sigma_l^2} - 2 \ln \frac{\sigma_l}{\sigma_m}. \end{aligned}$$

The quadratic formula can be applied to give two possible values for T_{lm} .

$$T_{lm} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}. \quad (A-6)$$

Normally only one of the two roots calculated from Eq. (A-6) will satisfy the condition $\mu_l < T_{lm} < \mu_m$. The root satisfying this condition is the proper threshold for a maximum Bayesian classifier separating classes l and m . Here all classes are assumed to be of equal *a priori* probability. The procedure would be the same, but the resulting expressions for a, b , and c would be more complex if different *a priori* probabilities were assigned to each class.

The performance of the maximum Bayesian classifier can be summarized by a confusion matrix, $C = c_{ij}$, where the i, j element is the probability of classifying a sample from class i into class j . If classification is perfect, C is the identity matrix. For imperfect classification, the location and magnitude of the nonzero, off-diagonal elements indicate the expected classification error rates for separation of each of the possible i, j class pairs.

Consider the special case of three classes of equal *a priori* probabilities where classes are numbered 1, 2, and 3 in order of increasing mean values of

x within the class. Thresholds T_{12} and T_{23} are then derived according to Eqs. (A-5) and (A-6). The confusion matrix for this special case can be calculated by

$$\begin{aligned} c_{i1} &= \int_{-\infty}^{T_{12}} P_i(x) dx \\ c_{i2} &= 1 - c_{i1} - c_{i3} \\ c_{i3} &= \int_{T_{23}}^{\infty} P_i(x) dx. \end{aligned} \quad (A-7)$$

Confusion matrices have the following mathematical properties.

$$\sum_i \sum_j c_{ij} = \text{number of classes} \quad (A-8)$$

$$\sum_j c_{ij} = 1 \quad (A-9)$$

the overall classification accuracy, A , can be calculated from the confusion matrix by the following equation:

$$A = \frac{\sum_i c_{ii}}{\sum_i \sum_j c_{ij}} = \frac{\text{Tr } C}{\text{number of classes}}. \quad (A-10)$$

Figure A-1 shows the normal distributions by ice type for the Laws' Mask 2, LM2, texture data from Table C-1. The figure also shows the interclass Gaussian decision thresholds defined by Eqs. (A-5) and (A-6). The calculated classification accuracy (Eq. (A-10)) for this case is 0.801.

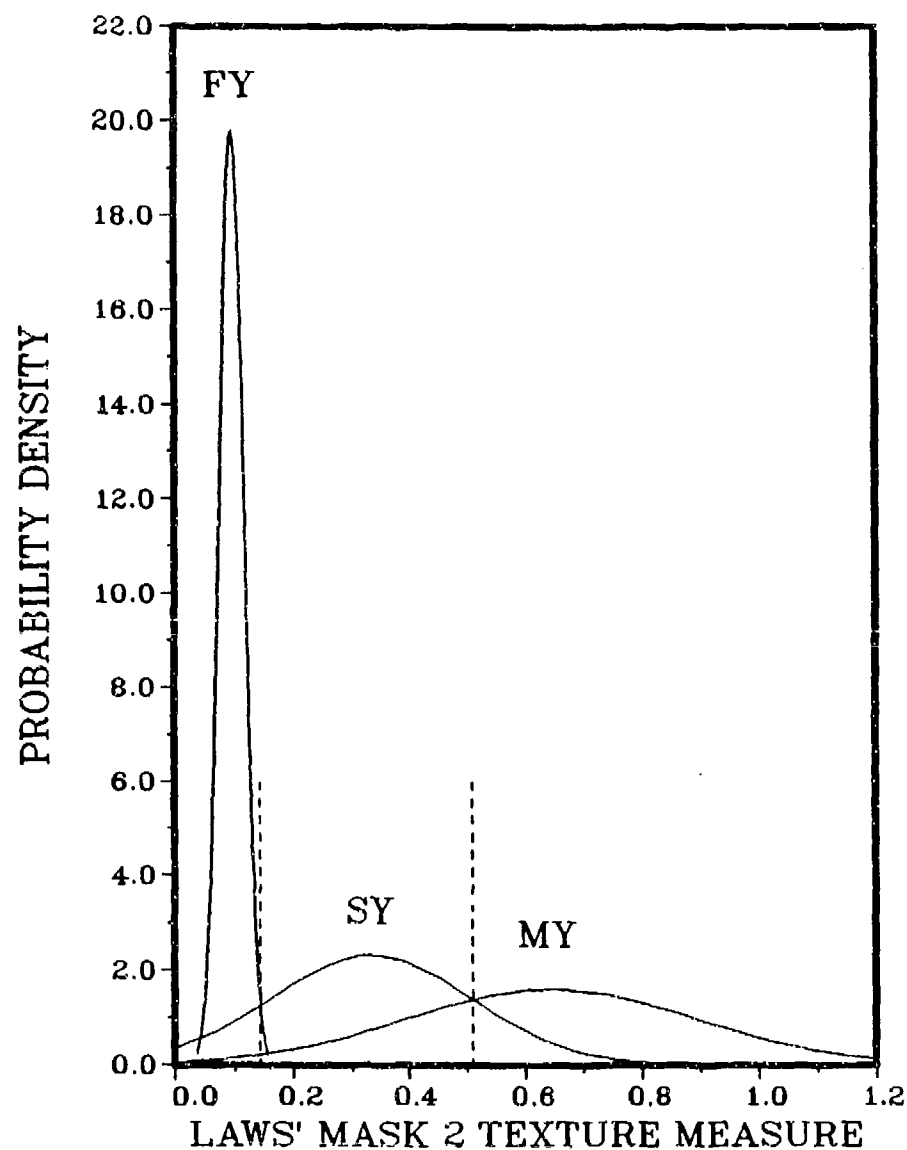


FIGURE A-1. Distributions of the Laws' Mask 2 texture measure values from data set A for first-year (FY), second-year (SY), and multiyear (MY) ice types. Dashed lines mark the maximum Bayesian decision boundaries between FY/SY and SY/MY classes. Classification accuracy for this texture measure is 0.801.

APPENDIX B

TESTING DATA SET NORMALITY: THE Q-Q PLOT

The Q-Q plot is a plot of the sample quantile versus the quantile one would expect from to observe if the observations are normally distributed. For details on Q-Q plots see Johnson and Wichern (1982). To summarize we list the following calculations leading to the Q-Q plot.

1. Arrange the original observations, X_i , to put X_1, X_2, \dots, X_n in order of increasing value.
2. Calculate P_1, P_2, \dots, P_n according to $P_i = \frac{(i-\frac{1}{2})}{n}$.
3. Calculate the standard normal quantiles q_1, q_2, \dots, q_n by integration of the normal distribution function

$$P_i = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{q_i} e^{-\frac{z^2}{2}} dz,$$

where $z = \frac{(X-\mu)}{\sigma}$.

4. Plot the data pairs (q_i, X_i) thereby producing the Q-Q plot.

If the values of X_i are normally distributed, the (q_i, X_i) points will fall in a straight line. "Straightness" of the Q-Q plot is therefore a useful measure of the normality of the distribution of X_i . The linear correlation coefficient of the (q_i, X_i) points is a good measure of straightness. Figure B-1 is given to illustrate the appearance of a Q-Q plot. This figure is for the Laws' Mask 2 texture measure, LM2, listed in Table C-1. Separate Q-Q plots are shown for the second-year and multiyear ice types. This example is typical of the texture measures considered in this study. The linear correlation coefficients for for second-year and multi-year ice types in Figure B-1 are 0.95 and 0.98, respectively.

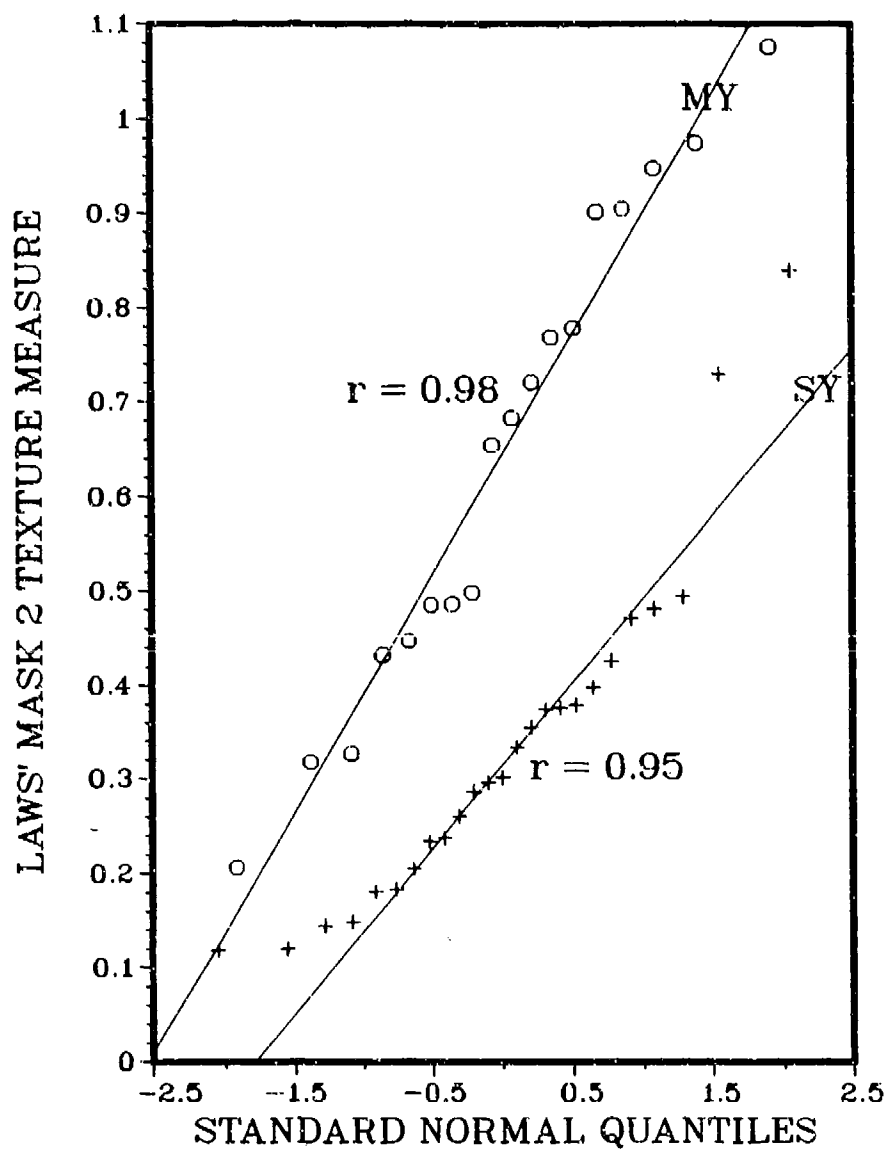


FIGURE B-1. Q-Q plots for the Laws' Mask 2 texture measure calculated from data set A. Second-year (SY) ice samples indicated by + and multiyear (MY) samples by o. Correlation coefficients for SY and MY ice are 0.95 and 0.98, respectively.

APPENDIX C

DATA SET A

Data set A consists of numerical values for each of the 25 candidate texture measures listed in Table III-1 for each of the fifty 64×64 pixel KRMS image samples identified in Figures IV-1 through IV-10. Table C-1 contains these numerical values. The values of the candidate texture measures varied by 23 orders of magnitude, so formatting of these data in tabular form was difficult. The texture variables were therefore scaled so that each contained only one digit to the left of the decimal point. In other words, each variable was scaled by some power of 10, so that its largest value in the data set would be less than 10 but greater than 1. The scaled data are given in Table C-1. However, scale factors required to convert the tabular data to their original values are included as the second column on each page of the table. For example, the MEAN variable for sample FY1 is given as 0.447. The scale factor (S.F.) corresponding to the MEAN is 10^2 . Thus, the true value of MEAN for sample FY1 is 0.447×10^2 , or 44.7.

Variables LPR3 and LPR15 in Table C-1 are logarithmically transformed versions of variables PR3 and PR15. See Section IV-B for a discussion of the reason for and exact nature of this transformation.

TABLE C-1. VALUES OF TEXTURE MEASURES FOR DATA SET A

MEAS.	S.F.	FY1	FY2	FY3	FY4	FY5	FY6	FY7	SY1
MEAN	10 ²	0.447	0.419	0.460	0.467	0.341	0.466	0.432	1.748
VAR	10 ²	0.474	0.391	0.372	0.314	0.333	0.360	0.406	0.699
SKEW	10 ⁰	3.170	3.350	3.012	2.951	3.008	3.253	3.096	2.723
KURT	10 ⁰	2.850	3.340	2.920	2.770	2.980	3.100	3.010	3.150
INR3	10 ²	0.453	0.475	0.481	0.430	0.417	0.437	0.462	0.600
INR15	10 ²	0.776	0.690	0.672	0.535	0.569	0.568	0.612	1.334
PR3	10 ⁶	0.060	0.043	0.030	0.019	0.025	0.029	0.042	0.155
PR15	10 ⁶	0.035	0.026	0.020	0.014	0.017	0.022	0.030	0.062
HOM3	10 ⁻¹	1.626	1.584	1.650	1.748	1.770	1.743	1.590	1.566
HOM15	10 ⁻¹	1.310	1.434	1.351	1.539	1.492	1.515	1.508	1.037
ENG3	10 ⁻²	2.215	2.570	2.530	2.825	2.855	2.715	2.430	1.675
ENG15	10 ⁻²	2.030	2.480	2.440	2.835	2.840	2.655	2.415	1.490
ENT3	10 ⁰	6.352	6.230	6.232	6.100	6.125	6.171	6.277	6.631
ENT15	10 ⁰	6.426	6.260	6.240	6.096	6.109	6.184	6.274	6.733
COR3	10 ⁻¹	5.258	3.887	3.643	3.196	3.772	3.950	4.352	5.732
COR15	10 ⁻¹	2.027	1.331	1.131	1.484	1.410	2.193	2.703	0.490
LM1	10 ²⁰	0.153	0.114	0.106	0.106	0.082	0.181	0.157	0.451
LM2	10 ¹⁶	0.125	0.092	0.118	0.064	0.088	0.082	0.112	0.376
LM3	10 ¹²	0.344	0.338	0.341	0.418	0.387	0.442	0.498	0.745
LM4	10 ¹³	0.564	0.646	0.601	0.695	0.675	0.798	0.857	0.761
LM5	10 ⁹	2.275	2.728	2.359	2.578	2.602	2.920	3.215	2.425
LM6	10 ⁸	1.759	2.075	1.735	1.765	1.700	1.878	2.223	1.573
LM7	10 ⁷	2.258	2.383	2.206	2.211	1.939	2.116	2.698	1.759
EDEN	10 ¹	1.911	1.931	1.818	1.732	1.744	1.735	2.015	2.197
MBRL	10 ⁰	3.626	3.321	3.480	3.208	3.691	3.511	3.490	4.759
LPR3	10 ⁰	2.187	1.853	1.493	1.037	1.311	1.460	1.830	3.136
LPR15	10 ⁰	1.648	1.350	1.088	0.731	0.925	1.183	1.493	2.219

TABLE C-1. continued

MEAS.	S.F.	SY2	SY3	SY4	SY5	SY6	SY7	SY8	SY9
MEAN	10 ²	1.751	1.734	1.778	1.795	1.784	1.720	1.761	1.744
VAR	10 ²	0.790	1.444	1.703	0.548	0.525	0.594	0.926	1.212
SKEW	10 ⁰	2.544	2.399	1.983	2.787	3.317	2.631	2.779	2.355
KURT	10 ⁰	3.530	3.240	3.950	3.020	3.250	3.160	2.490	3.350
INR3	10 ²	0.584	0.818	0.770	0.500	0.548	0.480	0.448	0.710
INR15	10 ²	1.380	2.710	2.525	0.949	0.989	1.162	0.934	1.904
PR3	10 ⁰	0.234	0.754	1.323	0.088	0.083	0.114	0.246	0.573
PR15	10 ⁰	0.116	0.265	0.547	0.050	0.035	0.044	0.159	0.212
HOM3	10 ⁻¹	1.498	1.362	1.444	1.628	1.563	1.696	1.690	1.366
HOM15	10 ⁻¹	1.054	0.714	0.867	1.189	1.154	1.147	1.177	0.911
ENG3	10 ⁻³	1.695	1.170	1.335	2.095	2.065	2.005	1.585	1.340
ENG15	10 ⁻³	1.415	0.940	1.065	1.860	1.885	1.700	1.460	1.195
ENT3	10 ⁰	6.653	7.000	6.918	6.472	6.448	6.471	6.360	6.879
ENT15	10 ⁰	6.796	7.163	7.081	6.537	6.507	6.620	6.735	6.965
COR3	10 ⁻¹	6.307	7.167	7.741	5.530	4.824	5.975	7.596	7.097
COR15	10 ⁻¹	1.721	0.677	2.950	1.460	0.461	0.418	5.404	1.636
LM1	10 ²⁰	0.449	1.000	1.901	0.311	0.198	0.388	0.665	0.603
LM2	10 ¹⁰	0.355	0.426	0.839	0.296	0.286	0.333	0.181	0.374
LM3	10 ¹³	0.716	0.828	1.248	0.602	0.577	0.646	0.491	0.734
LM4	10 ¹³	0.571	0.867	0.843	0.678	0.608	0.647	0.606	0.525
LM5	10 ⁹	1.621	2.723	2.187	2.154	1.972	1.849	1.911	1.804
LM6	10 ⁸	1.193	1.783	1.548	1.519	1.291	1.220	1.228	1.359
LM7	10 ⁷	1.657	2.300	1.984	1.781	1.594	1.513	1.369	1.922
EDEN	10 ¹	2.186	2.446	2.244	1.842	1.924	1.857	1.826	2.347
MBRL	10 ⁰	4.565	5.191	5.181	4.300	4.342	4.911	4.969	5.501
LPR3	10 ⁰	3.548	4.718	5.280	2.570	2.511	2.828	3.598	4.443
LPR15	10 ⁰	2.846	3.672	4.397	2.004	1.648	1.876	3.181	3.449

TABLE C-1. continued

MEAS.	S.F.	SY10	SY11	SY12	SY13	SY14	SY15	SY16	SY17
MEAN	10 ³	1.762	1.815	1.858	1.627	1.914	1.504	1.499	1.480
VAR	10 ³	1.261	1.346	0.458	1.586	0.385	0.538	0.504	0.458
SKEW	10 ⁰	2.904	1.687	2.467	2.250	2.741	2.723	2.857	2.937
KURT	10 ⁰	3.030	5.480	3.770	3.680	3.430	3.150	2.800	3.040
INR3	10 ³	0.865	0.604	0.495	0.758	0.442	0.510	0.494	0.506
INR15	10 ³	2.416	1.806	0.802	2.445	0.712	1.108	0.882	0.796
PR3	10 ⁶	0.562	1.288	0.076	1.067	0.043	0.085	0.062	0.049
PR15	10 ⁶	0.237	0.440	0.041	0.486	0.020	0.037	0.032	0.029
HOM3	10 ⁻¹	1.302	1.598	1.580	1.353	1.676	1.598	1.636	1.593
HOM15	10 ⁻¹	0.759	1.073	1.310	0.792	1.406	1.128	1.204	1.312
ENG3	10 ⁻³	1.160	1.625	2.335	1.145	2.655	2.055	2.050	2.220
ENG15	10 ⁻³	0.955	1.435	2.235	0.975	2.525	1.840	1.960	2.165
ENT3	10 ⁰	7.003	6.713	6.334	7.005	6.208	6.444	6.414	6.355
ENT15	10 ⁰	7.157	6.810	6.368	7.135	6.233	6.543	6.454	6.379
COR3	10 ⁻¹	6.630	7.815	4.793	7.603	4.324	5.259	5.092	4.449
COR15	10 ⁻¹	0.941	3.378	1.557	2.584	0.762	0.055	1.188	1.291
LM1	10 ²⁰	0.841	0.451	0.161	1.250	0.092	0.299	0.259	0.168
LM2	10 ¹⁶	0.729	0.144	0.119	0.481	0.118	0.238	0.148	0.183
LM3	10 ¹³	0.938	0.404	0.505	0.887	0.409	0.576	0.401	0.592
LM4	10 ¹³	0.811	0.614	0.680	0.761	0.530	0.695	0.598	0.788
LM5	10 ⁹	2.404	2.235	2.440	1.921	2.123	2.397	2.297	2.673
LM6	10 ⁸	1.508	1.534	1.766	1.276	1.694	1.807	1.701	1.960
LM7	10 ⁷	1.891	1.892	2.243	1.678	2.376	2.143	2.116	2.380
EDEN	10 ¹	2.508	2.018	1.854	2.267	1.835	1.858	1.843	1.878
MBRL	10 ⁰	4.958	5.020	3.632	6.812	3.440	4.011	4.157	4.142
LPR3	10 ⁰	4.424	5.253	2.423	5.065	1.853	2.535	2.219	1.984
LPR15	10 ⁰	3.560	4.179	1.806	4.278	1.088	1.703	1.558	1.460

TABLE C-1. continued

MEAS.	S.F.	SY18	SY19	SY20	SY21	SY22	SY23	SY24	SY25
MEAN	10 ²	1.484	1.623	1.634	1.666	1.715	1.586	1.634	1.607
VAR	10 ²	0.455	0.566	0.752	1.198	0.620	0.940	0.816	1.014
SKEW	10 ⁰	2.961	3.287	2.704	2.852	3.160	3.038	2.423	2.939
KURT	10 ⁰	2.950	3.730	3.280	3.400	3.350	4.360	3.870	3.060
INR3	10 ²	0.496	0.501	0.488	0.694	0.570	0.584	0.605	0.654
INR15	10 ²	0.847	1.080	1.167	2.207	1.197	1.690	1.584	2.211
PR3	10 ⁸	0.053	0.114	0.210	0.583	0.123	0.459	0.308	0.355
PR15	10 ⁸	0.027	0.045	0.113	0.236	0.053	0.147	0.083	0.103
HOM3	10 ⁻¹	1.604	1.639	1.615	1.421	1.488	1.516	1.515	1.466
HOM15	10 ⁻¹	1.287	1.183	1.106	0.800	1.077	0.993	1.019	0.782
ENG3	10 ⁻³	2.185	2.095	1.755	1.270	1.830	1.665	1.685	1.435
ENG15	10 ⁻³	2.060	1.900	1.470	1.000	1.655	1.385	1.525	1.130
ENT3	10 ⁰	6.370	6.436	6.607	6.911	6.546	6.689	6.666	6.824
ENT15	10 ⁰	6.412	6.527	6.756	7.107	6.636	6.837	6.749	6.992
COR3	10 ⁻¹	4.549	5.577	6.788	7.169	5.434	6.874	6.376	6.808
COR15	10 ⁻¹	0.772	0.298	2.722	1.326	0.414	0.928	0.000	0.000
LM1	10 ²⁰	0.105	0.201	0.417	1.104	0.308	0.556	0.295	0.799
LM2	10 ¹⁶	0.234	0.260	0.205	0.471	0.398	0.379	0.301	0.494
LM3	10 ¹³	0.632	0.554	0.441	0.909	0.634	0.552	0.592	0.682
LM4	10 ¹³	0.812	0.613	0.680	0.933	0.668	0.553	0.649	0.629
LM5	10 ⁹	2.260	2.045	2.455	2.682	2.271	1.885	2.439	2.012
LM6	10 ⁰	1.489	1.392	1.614	1.659	1.672	1.293	1.884	1.396
LM7	10 ⁷	1.964	1.548	1.763	1.839	2.091	1.529	2.486	1.824
EDEN	10 ¹	1.851	2.130	2.131	2.341	2.153	2.013	1.997	2.139
MBRL	10 ⁰	4.004	4.417	4.779	5.674	4.452	4.714	4.331	5.110
LPR3	10 ⁰	2.063	2.828	3.439	4.460	2.904	4.221	3.822	3.964
LPR15	10 ⁰	1.388	1.899	2.820	3.556	2.063	3.083	2.511	2.727

TABLE C-1. continued

MEAS.	S.F.	MY1	MY2	MY3	MY4	MY5	MY6	MY7	MY8
MEAN	10 ²	1.126	1.061	1.071	1.149	1.135	1.243	1.097	1.048
VAR	10 ²	1.473	1.438	1.498	1.177	1.139	1.626	2.660	2.416
SKEW	10 ⁰	2.172	2.466	2.569	2.193	2.609	2.761	2.540	2.855
KURT	10 ⁰	3.630	3.290	2.960	5.180	2.970	2.450	3.208	2.800
INR3	10 ²	0.927	0.757	0.735	0.822	0.872	0.967	1.296	1.273
INR15	10 ²	2.792	2.569	2.571	2.268	2.267	3.703	4.248	4.557
PR3	10 ⁶	0.857	0.848	0.808	0.793	0.380	0.745	2.742	1.984
PR15	10 ⁶	0.219	0.276	0.335	0.211	0.128	0.186	0.637	0.715
HOM3	10 ⁻¹	1.235	1.303	1.386	1.302	1.241	1.188	1.110	1.093
HOM15	10 ⁻¹	0.778	0.704	0.870	0.840	0.786	0.624	0.599	0.582
ENG3	10 ⁻³	1.155	1.115	1.100	1.220	1.150	0.325	0.745	0.730
ENG15	10 ⁻³	0.995	0.945	0.900	1.070	1.035	0.795	0.670	0.650
ENT3	10 ⁰	7.013	7.011	7.031	6.931	6.970	7.176	7.386	7.395
ENT15	10 ⁰	7.111	7.149	7.205	7.028	7.052	7.294	7.457	7.487
COR3	10 ⁻¹	6.862	7.399	7.549	6.550	6.151	7.029	7.550	7.386
COR15	10 ⁻¹	0.094	0.683	1.794	0.147	0.000	0.000	1.370	0.504
LM1	10 ²⁰	0.953	0.583	1.208	0.419	0.480	1.027	1.325	1.538
LM2	10 ¹⁶	0.768	0.485	0.486	0.447	0.682	0.901	1.075	0.974
LM3	10 ¹³	1.297	0.639	0.622	0.702	1.008	1.124	1.408	1.175
LM4	10 ¹²	0.979	0.654	0.611	0.792	0.834	0.955	0.990	1.387
LM5	10 ⁹	2.436	2.040	1.911	2.764	2.936	2.960	3.106	2.350
LM6	10 ⁸	1.525	1.351	1.357	1.640	2.189	1.923	2.288	1.498
LM7	10 ⁷	1.872	1.475	1.828	1.721	2.600	2.182	2.676	1.883
EDEN	10 ¹	2.584	2.486	2.498	2.539	2.580	2.518	2.932	2.986
MBRL	10 ⁰	5.686	6.312	6.084	4.831	5.147	6.817	6.159	7.082
LPR3	10 ⁰	4.846	4.835	4.787	4.768	4.032	4.706	6.009	5.685
LPR15	10 ⁰	3.481	3.713	3.906	3.444	2.944	3.318	4.822	4.665

TABLE C-1. continued

MEAS.	S.F.	MY9	MY10	MY11	MY12	MY13	MY14	MY15	MY16
MEAN	10 ²	0.993	1.429	1.425	1.394	1.608	1.538	1.567	1.532
VAR	10 ²	1.771	1.253	2.390	2.528	0.616	1.258	1.668	2.205
SKEW	10 ⁰	3.064	2.179	1.736	2.185	2.691	2.449	2.286	2.355
KURT	10 ⁰	2.750	3.920	5.470	4.450	2.930	3.330	3.120	3.800
INR3	10 ²	1.092	0.685	1.107	1.016	0.479	0.811	0.905	0.937
INR15	10 ²	3.067	2.080	3.248	3.594	0.931	2.120	2.962	3.643
PR3	10 ⁶	0.975	0.827	3.746	3.284	0.115	0.549	0.970	2.095
PR15	10 ⁶	0.397	0.404	1.444	1.023	0.069	0.264	0.446	0.607
HOM3	10 ⁻¹	1.120	1.521	1.245	1.629	1.632	1.417	1.288	1.275
HOM15	10 ⁻¹	0.687	0.921	0.722	0.731	1.213	0.843	0.713	0.641
ENG3	10 ⁻⁸	0.890	1.360	1.015	0.920	1.900	1.300	1.110	0.935
ENG15	10 ⁻⁸	0.790	1.130	0.895	0.770	1.725	1.025	0.900	0.790
ENT3	10 ⁰	7.228	6.865	7.136	7.216	6.505	6.920	7.061	7.196
ENT15	10 ⁰	7.308	7.016	7.220	7.358	6.588	7.091	7.213	7.330
COR3	10 ⁻¹	6.891	7.451	7.667	7.957	6.176	6.772	7.279	7.840
COR15	10 ⁻¹	1.263	2.120	2.594	2.551	2.670	1.794	1.496	1.079
LM1	10 ²⁰	1.151	0.698	1.704	1.441	0.280	1.078	1.627	1.736
LM2	10 ¹⁶	0.947	0.326	0.778	0.498	0.206	0.654	0.904	0.720
LM3	10 ¹⁸	0.834	0.602	1.181	0.827	0.448	1.677	1.606	1.242
LM4	10 ¹⁸	0.927	0.653	0.921	0.791	0.645	1.184	1.328	0.861
LM5	10 ⁹	2.546	2.063	2.398	2.805	2.631	2.706	3.203	2.103
LM6	10 ⁸	1.585	1.206	1.570	1.896	1.952	1.621	1.867	1.342
LM7	10 ⁷	1.726	1.213	1.939	2.181	2.351	2.019	2.047	1.654
EDEN	10 ¹	2.915	2.218	2.650	2.673	1.901	2.365	2.480	2.512
MBRL	10 ⁰	5.959	6.130	6.145	6.481	4.438	5.868	6.054	6.087
LPR3	10 ⁰	4.975	4.810	6.321	6.189	2.837	4.400	4.970	5.740
LPR15	10 ⁰	4.076	4.094	5.367	5.023	2.326	3.668	4.193	4.501

TABLE C-1. continued

MEAS.	S.F.	MY17	MY18
MEAN	10^2	1.436	1.466
VAR	10^2	1.025	1.214
SKEW	10^0	2.148	2.802
KURT	10^0	4.140	3.250
INR3	10^2	0.671	0.750
INR15	10^2	1.533	1.800
PR3	10^6	0.543	0.597
PR15	10^6	0.159	0.288
HOM3	10^{-1}	1.418	1.379
HOM15	10^{-1}	0.994	0.874
ENG3	10^{-2}	1.510	1.230
ENG15	10^{-3}	1.420	1.115
ENT3	10^0	6.772	6.955
ENT15	10^0	6.811	7.022
COR3	10^{-1}	6.749	7.029
COR15	10^{-1}	1.724	2.663
LM1	10^{20}	0.214	0.534
LM2	10^{16}	0.317	0.432
LM3	10^{13}	0.772	0.579
LM4	10^{13}	0.877	0.587
LM5	10^9	2.830	2.071
LM6	10^8	1.869	1.518
LM7	10^7	2.025	1.853
EDEN	10^1	2.222	2.279
MBRL	10^0	4.660	4.351
LPR3	10^0	4.389	4.484
LPR15	10^0	3.161	3.755

APPENDIX D

DATA SET B

Data set B consists of a subset of the candidate texture variables contained in data set A. Twelve of the original 25 texture variables in data set A were eliminated in the formation of data set B for reasons delineated in Section IV. MEAN has also been eliminated, since it is not a true texture measure. The 12 texture measures of data set B have been processed by the scale, invert, then force transformation described in Section V. This data matrix, therefore, contains row sums with a constant value of 1.0.

TABLE D-1. VALUES OF TEXTURE MEASURES FOR DATA SET B

	VAR	INR3	INR15	HOM3	HOM15	LM1	LM2	LM3	LM4	LM5	EDEN	MBRL
FY1	0.012	0.165	0.010	0.136	0.041	0.165	0.010	0.172	0.008	0.102	0.025	0.154
FY2	0.006	0.172	0.007	0.134	0.020	0.181	0.005	0.184	0.026	0.056	0.029	0.179
FY3	0.004	0.164	0.006	0.145	0.035	0.174	0.009	0.176	0.016	0.095	0.012	0.164
FY4	0.000	0.180	0.000	0.177	0.000	0.180	0.000	0.172	0.036	0.073	0.000	0.183
FY5	0.001	0.182	0.002	0.182	0.009	0.182	0.004	0.175	0.032	0.070	0.002	0.159
FY6	0.004	0.184	0.002	0.181	0.005	0.178	0.003	0.174	0.060	0.035	0.000	0.174
FY7	0.008	0.183	0.004	0.141	0.006	0.184	0.009	0.169	0.074	0.000	0.043	0.178
SY1	0.028	0.134	0.034	0.118	0.089	0.135	0.052	0.118	0.046	0.084	0.063	0.101
SY2	0.033	0.131	0.034	0.097	0.082	0.129	0.046	0.116	0.009	0.161	0.058	0.105
SY3	0.079	0.089	0.089	0.065	0.142	0.082	0.059	0.104	0.065	0.051	0.094	0.080
SY4	0.100	0.101	0.084	0.088	0.119	0.090	0.130	0.054	0.062	0.109	0.069	0.083
SY5	0.017	0.156	0.018	0.136	0.063	0.150	0.039	0.138	0.030	0.114	0.015	0.123
SY6	0.015	0.145	0.019	0.118	0.069	0.160	0.037	0.140	0.016	0.133	0.026	0.121
SY7	0.020	0.154	0.026	0.148	0.068	0.138	0.044	0.128	0.023	0.142	0.017	0.093
SY8	0.045	0.166	0.017	0.152	0.065	0.117	0.020	0.153	0.016	0.141	0.013	0.094
SY9	0.064	0.112	0.057	0.068	0.110	0.120	0.051	0.118	0.000	0.149	0.082	0.069
SY10	0.064	0.078	0.074	0.049	0.130	0.093	0.105	0.088	0.053	0.081	0.098	0.087
SY11	0.072	0.129	0.052	0.123	0.080	0.131	0.013	0.156	0.017	0.101	0.038	0.088
SY12	0.011	0.165	0.012	0.130	0.043	0.173	0.010	0.158	0.032	0.088	0.018	0.161
SY13	0.095	0.107	0.083	0.067	0.136	0.062	0.072	0.103	0.048	0.142	0.074	0.012
SY14	0.005	0.169	0.008	0.150	0.024	0.173	0.009	0.165	0.001	0.119	0.014	0.163
SY15	0.017	0.155	0.025	0.129	0.074	0.152	0.030	0.142	0.034	0.089	0.017	0.137
SY16	0.014	0.161	0.015	0.141	0.062	0.159	0.015	0.168	0.015	0.101	0.016	0.133
SY17	0.011	0.166	0.012	0.137	0.044	0.176	0.022	0.150	0.056	0.063	0.022	0.140
SY18	0.010	0.156	0.013	0.130	0.045	0.170	0.029	0.134	0.057	0.103	0.016	0.136
SY19	0.018	0.147	0.022	0.131	0.061	0.152	0.032	0.137	0.017	0.120	0.052	0.112
SY20	0.031	0.155	0.026	0.130	0.076	0.137	0.024	0.156	0.030	0.080	0.054	0.190
SY21	0.065	0.118	0.072	0.083	0.133	0.075	0.069	0.099	0.082	0.058	0.084	0.063
SY22	0.022	0.139	0.028	0.098	0.081	0.147	0.056	0.131	0.028	0.100	0.056	0.114
SY23	0.043	0.132	0.047	0.102	0.093	0.120	0.051	0.137	0.005	0.136	0.036	0.099
SY24	0.036	0.133	0.044	0.106	0.092	0.149	0.040	0.137	0.024	0.082	0.036	0.120
SY25	0.048	0.116	0.066	0.088	0.126	0.097	0.068	0.119	0.019	0.120	0.052	0.081
MY1	0.082	0.070	0.093	0.035	0.132	0.086	0.115	0.047	0.087	0.081	0.113	0.060
MY2	0.075	0.096	0.079	0.049	0.137	0.114	0.065	0.121	0.023	0.115	0.094	0.031
MY3	0.082	0.104	0.082	0.070	0.114	0.062	0.068	0.128	0.016	0.133	0.099	0.042
MY4	0.060	0.088	0.070	0.050	0.119	0.133	0.062	0.119	0.051	0.046	0.105	0.095
MY5	0.060	0.082	0.073	0.037	0.134	0.133	0.104	0.085	0.061	0.030	0.115	0.085
MY6	0.095	0.064	0.134	0.024	0.162	0.082	0.140	0.070	0.085	0.027	0.106	0.012
MY7	0.160	0.000	0.148	0.004	0.157	0.051	0.160	0.032	0.086	0.011	0.153	0.038
MY8	0.129	0.004	0.144	0.000	0.144	0.029	0.130	0.054	0.144	0.078	0.144	0.000
MY9	0.096	0.036	0.098	0.006	0.138	0.064	0.135	0.098	0.072	0.065	0.146	0.045
MY10	0.067	0.116	0.064	0.106	0.108	0.110	0.043	0.134	0.025	0.121	0.065	0.041

TABLE D-1. continued

	VAR	INR3	INR15	HOM3	HOM15	LM1	LM2	LM3	LM4	LM5	EDEN	MBRL
MY11	0.148	0.036	0.113	0.038	0.143	0.018	0.118	0.062	0.077	0.086	0.122	0.040
MY12	0.146	0.049	0.118	0.123	0.131	0.039	0.067	0.098	0.048	0.040	0.116	0.024
MY13	0.023	0.167	0.018	0.143	0.061	0.160	0.025	0.165	0.025	0.066	0.024	0.123
MY14	0.073	0.100	0.072	0.087	0.132	0.082	0.106	0.000	0.139	0.058	0.092	0.057
MY15	0.103	0.079	0.108	0.051	0.154	0.027	0.148	0.009	0.166	0.001	0.106	0.047
MY16	0.129	0.066	0.124	0.043	0.151	0.015	0.104	0.052	0.063	0.112	0.100	0.041
MY17	0.052	0.122	0.043	0.082	0.098	0.159	0.043	0.116	0.070	0.041	0.067	0.107
MY18	0.061	0.099	0.050	0.067	0.110	0.119	0.058	0.130	0.011	0.114	0.069	0.112

APPENDIX E

DATA SET C

Data set C consists of 10 of the texture variables from data set B. LM4 and LM5 from B are not present in C. In addition to the 10 texture measures, this data set includes the mean image brightness value, MEAN. The 11 variables in data set C have been processed by the scale, invert, then force transformation described in Section V. This data matrix, therefore, contains row sums with a constant value of 1.0.

TABLE E-1. DATA SET C

	VAR	INR3	INR15	HOM3	HOM15	LM1	LM2	LM3	EDEN	MBRL	MEAN
FY1	0.013	0.183	0.011	0.151	0.046	0.183	0.012	0.190	0.027	0.171	0.013
FY2	0.007	0.186	0.008	0.144	0.022	0.195	0.006	0.199	0.032	0.193	0.010
FY3	0.005	0.181	0.007	0.161	0.038	0.193	0.010	0.195	0.013	0.182	0.015
FY4	0.000	0.199	0.000	0.195	0.000	0.199	0.000	0.190	0.000	0.202	0.016
FY5	0.002	0.203	0.002	0.203	0.010	0.203	0.005	0.195	0.002	0.177	0.000
FY6	0.004	0.200	0.002	0.197	0.005	0.194	0.004	0.189	0.000	0.189	0.016
FY7	0.008	0.195	0.004	0.151	0.007	0.197	0.010	0.181	0.046	0.190	0.012
SY1	0.027	0.131	0.033	0.116	0.087	0.132	0.051	0.115	0.061	0.099	0.148
SY2	0.034	0.134	0.035	0.099	0.084	0.132	0.048	0.119	0.060	0.108	0.148
SY3	0.077	0.087	0.086	0.064	0.138	0.079	0.057	0.101	0.091	0.078	0.142
SY4	0.102	0.103	0.085	0.089	0.121	0.000	0.132	0.055	0.070	0.085	0.157
SY5	0.017	0.153	0.017	0.134	0.062	0.148	0.039	0.136	0.015	0.122	0.157
SY6	0.015	0.144	0.019	0.118	0.068	0.159	0.037	0.139	0.026	0.120	0.155
SY7	0.020	0.157	0.026	0.151	0.069	0.141	0.045	0.130	0.017	0.095	0.148
SY8	0.045	0.167	0.017	0.152	0.065	0.117	0.020	0.153	0.013	0.094	0.156
SY9	0.064	0.112	0.057	0.068	0.110	0.120	0.051	0.118	0.082	0.068	0.150
SY10	0.064	0.077	0.074	0.049	0.128	0.092	0.104	0.087	0.097	0.086	0.142
SY11	0.070	0.125	0.050	0.118	0.077	0.127	0.013	0.151	0.036	0.084	0.149
SY12	0.011	0.156	0.011	0.123	0.041	0.164	0.009	0.150	0.017	0.153	0.165
SY13	0.099	0.112	0.087	0.070	0.143	0.065	0.075	0.108	0.078	0.013	0.150
SY14	0.005	0.160	0.007	0.142	0.023	0.164	0.009	0.156	0.014	0.155	0.165
SY15	0.016	0.154	0.024	0.128	0.074	0.151	0.030	0.141	0.017	0.136	0.127
SY16	0.014	0.159	0.015	0.139	0.061	0.157	0.014	0.166	0.015	0.131	0.128
SY17	0.011	0.164	0.012	0.135	0.043	0.174	0.021	0.148	0.021	0.138	0.132
SY18	0.011	0.162	0.014	0.134	0.047	0.176	0.030	0.139	0.017	0.141	0.129
SY19	0.018	0.148	0.022	0.132	0.061	0.153	0.032	0.137	0.052	0.113	0.133
SY20	0.031	0.151	0.026	0.126	0.074	0.134	0.023	0.151	0.052	0.097	0.135
SY21	0.065	0.117	0.071	0.083	0.130	0.075	0.069	0.098	0.083	0.062	0.144
SY22	0.022	0.136	0.027	0.096	0.080	0.145	0.055	0.129	0.055	0.112	0.144
SY23	0.044	0.133	0.047	0.103	0.094	0.122	0.051	0.138	0.037	0.101	0.130
SY24	0.035	0.129	0.043	0.102	0.089	0.145	0.038	0.133	0.035	0.116	0.135
SY25	0.048	0.118	0.067	0.089	0.128	0.098	0.069	0.120	0.052	0.082	0.130
MY1	0.089	0.076	0.102	0.038	0.144	0.094	0.126	0.051	0.123	0.065	0.090
MY2	0.080	0.103	0.085	0.052	0.147	0.122	0.000	0.130	0.101	0.033	0.077
MY3	0.089	0.112	0.089	0.076	0.123	0.067	0.000	0.138	0.107	0.045	0.081
MY4	0.061	0.089	0.071	0.051	0.121	0.135	0.060	0.121	0.107	0.096	0.085
MY5	0.060	0.083	0.074	0.037	0.135	0.134	0.105	0.086	0.116	0.085	0.086
MY6	0.096	0.064	0.136	0.024	0.165	0.083	0.143	0.071	0.108	0.012	0.099
MY7	0.163	0.000	0.151	0.004	0.160	0.052	0.163	0.033	0.156	0.039	0.078
MY8	0.153	0.004	0.171	0.000	0.171	0.034	0.154	0.064	0.171	0.000	0.077
MY9	0.104	0.039	0.105	0.007	0.149	0.069	0.146	0.105	0.158	0.049	0.069
MY10	0.069	0.120	0.066	0.109	0.111	0.114	0.045	0.138	0.067	0.042	0.119

TABLE E-1. continued

	VAR	INR3	INR15	HOM3	HOM15	LM1	LM2	LM3	EDEN	MBRL	MEAN
MY11	0.155	0.038	0.118	0.039	0.150	0.019	0.124	0.065	0.128	0.642	0.121
MY12	0.144	0.049	0.116	0.121	0.129	0.039	0.066	0.097	0.115	0.024	0.102
MY13	0.022	0.158	0.017	0.135	0.058	0.152	0.024	0.156	0.023	0.116	0.137
MY14	0.078	0.107	0.076	0.093	0.141	0.088	0.113	0.000	0.098	0.061	0.147
MY15	0.103	0.082	0.111	0.053	0.158	0.028	0.152	0.010	0.109	0.049	0.143
MY16	0.137	0.069	0.131	0.046	0.159	0.015	0.110	0.055	0.106	0.044	0.128
MY17	0.052	0.121	0.042	0.082	0.097	0.158	0.043	0.115	0.066	0.106	0.118
MY18	0.062	0.100	0.050	0.068	0.112	0.121	0.058	0.132	0.070	0.113	0.115

APPENDIX F

DATA SET D

Data set D consists of 42 synthetic mixture images. Table F-1 contains numerical values for mean intensity and for each of the 12 selected texture measures (VAR, INR3, INR15, HOM3, HOM15, LM1, LM2, LM3, LM4, LM5, EDEN, and MBRL) for each sample in data set D. The mixtures are labeled MX1 through MX42, and are 128×128 pixels in size. The KRMS synthetic mixture images are shown in Figures VII-1 through VII-7. Formatting and scaling of variables in this table is the same as in Table C-1.

Table F-1 also contains the compositions of each mixture by ice type. These are given at the bottom of the table.

TABLE F-1. VALUES OF TEXTURE MEASURES FOR DATA SET D

MEAS.	S.F.	MX1	MX2	MX3	MX4	MX5	MX6	MX7	MX8
MEAN	10 ²	1.200	1.286	1.355	1.728	1.736	1.656	0.418	0.413
VAR	10 ²	1.706	1.797	1.836	1.619	0.671	0.651	0.392	0.378
INR3	10 ²	1.021	1.085	1.068	0.670	0.625	0.545	0.453	0.468
INR15	10 ²	2.969	3.014	3.042	1.649	1.254	1.163	0.676	0.653
HOM3	10 ⁻¹	1.278	1.233	1.244	1.464	1.520	1.573	1.672	1.636
HOM15	10 ⁻¹	0.779	0.732	0.733	0.980	1.136	1.138	1.408	1.415
LM1	10 ²⁰	0.903	1.163	1.163	0.619	0.339	0.354	0.148	0.148
LM2	10 ¹⁶	0.857	0.843	0.766	0.403	0.345	0.247	0.103	0.104
LM3	10 ¹⁸	1.306	1.292	1.090	0.786	0.657	0.559	0.399	0.399
LM4	10 ¹⁸	1.041	1.016	0.928	0.707	0.710	0.691	0.685	0.685
LM5	10 ⁹	2.858	2.768	2.779	2.123	2.265	2.507	2.621	2.658
EDEN	10 ¹	2.648	2.722	2.717	2.194	2.078	1.992	1.842	1.897
MBRL	10 ⁰	6.143	6.482	6.115	5.661	4.325	4.604	3.673	3.722
% FY		0	0	0	0	0	0	100	100
% SY		0	0	0	100	100	100	0	0
% MY		100	100	100	0	0	0	0	0

TABLE F-1. continued

MEAS.	S.F.	MX9	MX10	MX11	MX12	MX13	MX14	MX15	MX16
MEAN	10 ²	0.427	0.931	0.939	1.332	1.323	1.190	1.277	0.783
VAR	10 ²	0.356	0.928	0.830	0.570	1.042	1.101	1.534	0.887
INR3	10 ²	0.442	0.632	0.642	0.525	0.659	0.704	0.843	0.640
INR15	10 ²	0.592	1.607	1.646	0.997	1.974	1.867	2.402	1.530
HOM3	10 ⁻¹	1.700	1.516	1.511	1.601	1.490	1.617	1.438	1.507
HOM15	10 ⁻¹	1.493	1.094	1.107	1.222	0.993	0.982	0.979	1.123
LM1	10 ²⁰	0.133	0.551	0.487	0.291	0.733	0.633	1.019	0.434
LM2	10 ¹⁶	0.092	0.275	0.348	0.225	0.409	0.372	0.456	0.331
LM3	10 ¹⁸	0.433	0.529	0.651	0.575	0.757	0.640	0.783	0.550
LM4	10 ¹⁸	0.739	0.650	0.737	0.703	0.765	0.715	0.770	0.655
LM5	10 ⁹	2.742	2.218	2.408	2.367	2.337	2.339	2.525	2.277
EDEN	10 ¹	1.818	2.169	2.060	1.961	2.168	2.239	2.337	2.202
MBRL	10 ⁰	3.617	4.384	4.449	4.155	5.059	5.265	5.373	4.330
% FY		100	50	50	25	25	25	25	50
% SY		0	25	25	50	50	25	25	0
% MY		0	25	25	25	25	50	50	50

TABLE F-1. continued

MEAS.	S.F.	MX17	MX18	MX19	MX20	MX21	MX22	MX23	MX24
MEAN	10^2	1.446	1.010	1.387	0.785	0.728	1.029	1.556	1.445
VAR	10^2	1.550	0.614	0.691	0.362	0.709	1.438	1.206	1.271
INR3	10^2	0.896	0.543	0.569	0.448	0.572	0.879	0.829	0.846
INR15	10^2	2.558	1.093	1.204	0.645	1.207	2.636	2.350	2.091
HOM3	10^{-1}	1.402	1.554	1.545	1.683	1.553	1.329	1.348	1.352
HOM15	10^{-1}	0.900	1.222	1.161	1.445	1.252	0.873	0.840	0.895
LM1	10^{20}	1.049	0.353	0.381	0.104	0.442	0.933	0.736	0.711
LM2	10^{16}	0.711	0.236	0.274	0.095	0.264	0.611	0.562	0.548
LM3	10^{13}	1.171	0.605	0.590	0.406	0.620	1.036	0.935	0.893
LM4	10^{13}	0.971	0.788	0.652	0.678	0.789	0.896	0.896	0.795
LM5	10^9	2.823	2.639	2.163	2.606	2.615	2.542	2.561	2.403
EDEN	10^1	2.395	2.022	2.107	1.824	2.044	2.471	2.383	2.402
MBRL	10^0	6.116	4.142	4.511	3.531	3.992	5.277	5.311	5.401
% FY		0	50	25	75	75	25	0	0
% SY		50	50	75	25	0	0	75	25
% MT		50	0	0	0	25	75	25	75

TABLE F-1. continued

MEAS.	S.F.	MX25	MX26	MX27	MX28	MX29	MX30	MX31	MX32
MEAN	10^2	1.077	1.214	1.105	1.004	0.894	1.239	0.830	0.883
VAR	10^2	1.008	1.010	0.984	1.041	0.512	0.694	0.653	1.095
INR3	10^2	0.689	0.655	0.697	0.792	0.493	0.508	0.528	0.687
INR15	10^2	1.831	1.787	1.781	1.995	0.818	1.117	0.922	1.812
HOM3	10^{-1}	1.445	1.494	1.415	1.457	1.629	1.597	1.607	1.418
HOM15	10^{-1}	0.969	1.045	0.961	1.083	1.284	1.162	1.336	1.007
LM1	10^{20}	0.599	0.734	0.567	0.541	0.208	0.400	0.417	0.702
LM2	10^{16}	0.343	0.383	0.411	0.484	0.151	0.250	0.197	0.398
LM3	10^{13}	0.556	0.776	0.689	0.776	0.428	0.499	0.560	0.595
LM4	10^{13}	0.613	0.847	0.728	0.780	0.601	0.641	0.724	0.655
LM5	10^9	2.118	2.572	2.364	2.539	2.172	2.227	2.516	2.351
EDEN	10^1	2.315	2.175	2.298	2.253	1.892	2.097	1.925	2.306
MBRL	10^0	4.940	4.658	4.954	4.659	4.025	4.717	4.103	5.095
% FY		33	33	33	33	67	33	67	33
% SY		33	33	33	33	33	67	0	0
% MY		33	33	33	33	0	0	33	67

TABLE F-1. continued

MEAS.	S.F.	MX33	MX34	MX35	MX36	MX37	MX38	MX39	MX40
MEAN	10^2	1.646	1.579	0.448	1.753	1.765	1.795	1.492	1.660
VAR	10^2	1.046	1.116	0.424	1.176	0.732	1.296	0.499	0.914
INR3	10^2	0.632	0.657	0.459	0.691	0.495	0.676	0.501	0.562
INR15	10^2	1.513	1.653	0.657	2.013	1.020	1.833	0.906	1.416
HOM3	10^{-1}	1.522	1.525	1.652	1.468	1.640	1.461	1.608	1.541
HOM15	10^{-1}	1.039	1.045	1.417	0.911	1.167	0.989	1.230	1.034
LM1	10^{20}	0.665	0.674	0.168	0.949	0.525	1.017	0.264	0.634
LM2	10^{16}	0.360	0.377	0.103	0.485	0.280	0.414	0.204	0.336
LM3	10^{13}	0.709	0.859	0.368	0.845	0.582	0.734	0.547	0.638
LM4	10^{13}	0.706	0.851	0.627	0.729	0.652	0.693	0.753	0.723
LM5	10^9	2.173	2.511	2.466	2.180	2.097	2.171	2.554	2.315
EDEN	10^1	2.083	2.215	1.847	2.266	1.866	2.186	1.859	2.186
MBRL	10^0	5.608	5.646	3.731	5.392	5.358	6.073	4.363	5.566
% FY		0	0	100	0	0	0	0	0
% SY		67	33	0	100	100	100	100	100
% MY		33	67	0	0	0	0	0	0

TABLE F-1. continued

MEAS.	S.F.	MX41	MX42
MEAN	10^2	1.561	1.104
VAR	10^2	1.498	1.486
INR3	10^2	0.778	0.794
INR15	10^2	2.429	2.427
HOM3	10^{-1}	1.405	1.310
HOM15	10^{-1}	0.863	0.791
LM1	10^{20}	1.269	0.886
LM2	10^{16}	0.557	0.525
LM3	10^{13}	1.144	0.687
LM4	10^{13}	0.986	0.674
LM5	10^9	2.665	2.333
EDEN	10^1	2.303	2.524
MBRL	10^0	6.571	6.340
% FY		0	0
% SY		0	0
% MY		100	100

APPENDIX G

DATA SET E

Data set E consists of a subset (the 12 texture measures) of the variables contained in data set D. MEAN is not included in E since it is not a true texture measure. The 12 texture measures of data set E have been processed by the scale, invert, then force transformation described in Section V. This data matrix, therefore, contains row sums with a constant value of 1.0.

TABLE G-1. DATA SET E

	VAR	INR3	INR15	HOM3	HOM15	LM1	LM2	LM3	LM4	LM5	EDEN	MBRL
MX1	0.143	0.016	0.152	0.015	0.147	0.049	0.157	0.000	0.157	0.000	0.144	0.022
MX2	0.159	0.000	0.161	0.000	0.163	0.015	0.160	0.002	0.154	0.019	0.163	0.005
MX3	0.161	0.004	0.161	0.004	0.160	0.012	0.142	0.037	0.119	0.017	0.160	0.024
MX4	0.073	0.105	0.070	0.081	0.110	0.091	0.066	0.090	0.039	0.157	0.068	0.049
MX5	0.035	0.116	0.044	0.100	0.076	0.130	0.054	0.112	0.040	0.127	0.047	0.120
MX6	0.035	0.146	0.040	0.126	0.081	0.136	0.035	0.138	0.036	0.080	0.033	0.112
MX7	0.004	0.178	0.006	0.170	0.020	0.174	0.003	0.175	0.035	0.056	0.005	0.173
MX8	0.003	0.178	0.005	0.160	0.019	0.179	0.003	0.179	0.035	0.049	0.016	0.174
MX9	0.000	0.187	0.000	0.187	0.000	0.182	0.000	0.174	0.059	0.029	0.000	0.182
MX10	0.061	0.110	0.065	0.095	0.082	0.097	0.037	0.130	0.017	0.132	0.061	0.113
MX11	0.052	0.113	0.070	0.097	0.083	0.110	0.055	0.114	0.051	0.097	0.044	0.114
MX12	0.024	0.146	0.028	0.132	0.060	0.141	0.029	0.131	0.039	0.108	0.027	0.134
MX13	0.074	0.105	0.090	0.087	0.104	0.073	0.066	0.093	0.059	0.109	0.061	0.079
MX14	0.077	0.090	0.079	0.125	0.102	0.083	0.056	0.108	0.039	0.104	0.071	0.065
MX15	0.131	0.062	0.122	0.072	0.111	0.035	0.078	0.092	0.063	0.072	0.095	0.065
MX16	0.056	0.108	0.060	0.092	0.076	0.112	0.049	0.126	0.019	0.119	0.066	0.115
MX17	0.138	0.050	0.137	0.062	0.133	0.032	0.138	0.025	0.143	0.008	0.109	0.026
MX18	0.030	0.147	0.036	0.120	0.062	0.137	0.033	0.131	0.074	0.050	0.039	0.140
MX19	0.037	0.130	0.040	0.108	0.071	0.123	0.039	0.124	0.019	0.148	0.052	0.110
MX20	0.001	0.179	0.004	0.175	0.011	0.181	0.001	0.174	0.032	0.060	0.001	0.181
MX21	0.041	0.138	0.043	0.118	0.055	0.122	0.039	0.126	0.074	0.055	0.043	0.146
MX22	0.114	0.050	0.130	0.032	0.127	0.045	0.106	0.045	0.105	0.065	0.113	0.057
MX23	0.090	0.063	0.113	0.039	0.135	0.072	0.097	0.062	0.105	0.061	0.098	0.065
MX24	0.099	0.060	0.098	0.041	0.126	0.077	0.096	0.071	0.071	0.096	0.104	0.062
MX25	0.068	0.095	0.078	0.070	0.106	0.089	0.051	0.123	0.004	0.150	0.085	0.083
MX26	0.072	0.109	0.080	0.091	0.096	0.075	0.062	0.092	0.092	0.062	0.065	0.103
MX27	0.068	0.096	0.077	0.062	0.111	0.096	0.066	0.105	0.045	0.103	0.085	0.085
MX28	0.075	0.074	0.093	0.078	0.088	0.102	0.083	0.092	0.066	0.068	0.078	0.102
MX29	0.018	0.154	0.015	0.142	0.046	0.152	0.013	0.156	0.000	0.151	0.014	0.140
MX30	0.037	0.145	0.035	0.126	0.070	0.120	0.033	0.139	0.015	0.134	0.050	0.098
MX31	0.036	0.157	0.024	0.145	0.037	0.132	0.025	0.144	0.051	0.081	0.021	0.147
MX32	0.082	0.101	0.081	0.065	0.105	0.080	0.035	0.124	0.020	0.109	0.088	0.079
MX33	0.077	0.117	0.062	0.103	0.089	0.086	0.058	0.106	0.040	0.150	0.049	0.053
MX34	0.086	0.112	0.073	0.105	0.089	0.086	0.063	0.080	0.095	0.077	0.074	0.051
MX35	0.008	0.176	0.005	0.162	0.018	0.170	0.003	0.180	0.011	0.093	0.006	0.169
MX36	0.087	0.096	0.091	0.079	0.120	0.043	0.081	0.077	0.046	0.140	0.078	0.061
MX37	0.043	0.156	0.030	0.148	0.073	0.109	0.042	0.131	0.020	0.170	0.009	0.068
MX38	0.108	0.109	0.086	0.083	0.113	0.037	0.072	0.104	0.036	0.154	0.069	0.028
MX39	0.017	0.162	0.023	0.143	0.062	0.154	0.026	0.144	0.062	0.069	0.008	0.130
MX40	0.062	0.133	0.055	0.108	0.099	0.089	0.052	0.117	0.045	0.117	0.067	0.054
MX41	0.137	0.085	0.133	0.065	0.147	0.060	0.108	0.031	0.155	0.045	0.095	0.000
MX42	0.121	0.072	0.119	0.026	0.146	0.052	0.090	0.104	0.026	0.109	0.124	0.012

APPENDIX H

DATA SET F

Data set F consists of 10 of the texture variables from data set E. LM4 and LM5 from E have are not present in F. In addition to the 10 texture measures, this data set includes the mean image brightness value, MEAN. The 11 variables in data set F have been processed by the scale, invert, then force transformation described in Section V. This data matrix, therefore, contains row sums with a constant value of 1.0.

TABLE H-1. DATA SET F

	VAR	INR3	INR15	HOM3	HOM15	LM1	LM2	LM3	EDEN	MBRL	MEAN
MX01	0.153	0.017	0.163	0.016	0.157	0.053	0.168	0.000	0.154	0.024	0.096
MX02	0.170	0.000	0.173	0.000	0.175	0.016	0.172	0.003	0.175	0.005	0.111
MX03	0.165	0.004	0.165	0.004	0.165	0.012	0.145	0.038	0.164	0.025	0.112
MX04	0.076	0.110	0.073	0.084	0.115	0.095	0.069	0.094	0.071	0.051	0.162
MX05	0.035	0.118	0.044	0.101	0.077	0.131	0.054	0.114	0.047	0.121	0.157
MX06	0.033	0.140	0.039	0.122	0.078	0.131	0.034	0.133	0.032	0.108	0.150
MX07	0.005	0.196	0.007	0.187	0.022	0.192	0.003	0.193	0.005	0.190	0.001
MX08	0.003	0.194	0.005	0.175	0.021	0.195	0.003	0.196	0.018	0.190	0.000
MX09	0.000	0.205	0.000	0.205	0.000	0.200	0.000	0.190	0.000	0.199	0.002
MX10	0.067	0.121	0.071	0.104	0.090	0.106	0.041	0.143	0.067	0.124	0.065
MX11	0.057	0.123	0.077	0.106	0.091	0.120	0.060	0.125	0.048	0.125	0.068
MX12	0.025	0.152	0.029	0.137	0.062	0.146	0.030	0.136	0.028	0.139	0.116
MX13	0.079	0.112	0.096	0.093	0.111	0.078	0.070	0.099	0.066	0.084	0.112
MX14	0.081	0.096	0.084	0.133	0.108	0.088	0.059	0.115	0.075	0.069	0.091
MX15	0.136	0.064	0.126	0.075	0.115	0.037	0.081	0.095	0.098	0.067	0.107
MX16	0.062	0.120	0.066	0.102	0.084	0.124	0.054	0.140	0.074	0.128	0.046
MX17	0.141	0.051	0.140	0.063	0.136	0.033	0.141	0.025	0.112	0.026	0.131
MX18	0.032	0.155	0.036	0.126	0.065	0.144	0.035	0.137	0.041	0.147	0.079
MX19	0.039	0.137	0.043	0.114	0.075	0.130	0.041	0.131	0.055	0.116	0.121
MX20	0.001	0.188	0.004	0.182	0.012	0.189	0.001	0.182	0.001	0.189	0.051
MX21	0.045	0.151	0.048	0.130	0.060	0.134	0.043	0.138	0.047	0.161	0.043
MX22	0.127	0.056	0.145	0.036	0.142	0.050	0.118	0.050	0.126	0.074	0.077
MX23	0.094	0.065	0.117	0.040	0.140	0.075	0.100	0.065	0.102	0.068	0.135
MX24	0.104	0.063	0.103	0.043	0.132	0.081	0.100	0.074	0.109	0.065	0.126
MX25	0.074	0.103	0.085	0.076	0.115	0.096	0.055	0.134	0.092	0.090	0.080
MX26	0.077	0.116	0.085	0.097	0.102	0.080	0.066	0.098	0.069	0.109	0.101
MX27	0.073	0.103	0.083	0.067	0.120	0.103	0.071	0.113	0.091	0.091	0.086
MX28	0.080	0.079	0.100	0.083	0.094	0.109	0.089	0.098	0.084	0.109	0.074
MX29	0.019	0.169	0.017	0.156	0.051	0.168	0.014	0.172	0.015	0.154	0.064
MX30	0.039	0.153	0.036	0.132	0.074	0.127	0.035	0.146	0.052	0.104	0.102
MX31	0.039	0.170	0.026	0.157	0.040	0.143	0.027	0.156	0.023	0.159	0.059
MX32	0.088	0.109	0.088	0.070	0.113	0.086	0.071	0.134	0.095	0.086	0.060
MX33	0.081	0.122	0.065	0.107	0.103	0.090	0.061	0.110	0.051	0.055	0.155
MX34	0.089	0.115	0.075	0.108	0.102	0.088	0.065	0.083	0.076	0.053	0.146
MX35	0.009	0.195	0.005	0.180	0.020	0.189	0.003	0.200	0.006	0.187	0.005
MX36	0.090	0.100	0.094	0.082	0.124	0.045	0.084	0.080	0.081	0.063	0.158
MX37	0.044	0.160	0.030	0.152	0.075	0.111	0.043	0.135	0.009	0.070	0.171
MX38	0.111	0.111	0.088	0.085	0.115	0.038	0.073	0.106	0.071	0.029	0.174
MX39	0.017	0.161	0.023	0.142	0.061	0.153	0.026	0.143	0.008	0.128	0.138
MX40	0.063	0.135	0.056	0.110	0.100	0.091	0.053	0.119	0.068	0.055	0.150
MX41	0.144	0.089	0.140	0.059	0.155	0.000	0.114	0.032	0.100	0.000	0.155
MX42	0.128	0.076	0.126	0.028	0.155	0.055	0.095	0.111	0.131	0.013	0.084

APPENDIX I

DATA SET G

Data set G is a subset of data set D. G contains only the texture variables from D that were shown in Section IV to mix linearly, even when the mean intensities of the two image halves were not equal. The six texture measures in Table I-1 (HOM3, HOM15, LM2, LM3, LM4, and LM5) have been SIFTed to row sums of 1.0.

TABLE I-1. DATA SET G

	HOM3	HOM15	LM2	LM3	LM4	LM5
MX01	0.032	0.309	0.000	0.330	0.000	0.330
MX02	0.000	0.340	0.006	0.335	0.019	0.300
MX03	0.008	0.326	0.039	0.251	0.084	0.292
MX04	0.165	0.225	0.198	0.148	0.253	0.011
MX05	0.203	0.155	0.221	0.112	0.248	0.073
MX06	0.206	0.132	0.226	0.058	0.225	0.153
MX07	0.263	0.031	0.276	0.009	0.227	0.193
MX08	0.245	0.029	0.279	0.009	0.229	0.209
MX09	0.278	0.000	0.278	0.019	0.190	0.235
MX10	0.195	0.169	0.245	0.055	0.286	0.051
MX11	0.188	0.160	0.210	0.095	0.218	0.129
MX12	0.238	0.107	0.249	0.067	0.232	0.107
MX13	0.175	0.209	0.186	0.132	0.199	0.100
MX14	0.237	0.193	0.182	0.083	0.213	0.091
MX15	0.135	0.207	0.161	0.136	0.189	0.173
MX16	0.191	0.158	0.224	0.063	0.286	0.077
MX17	0.110	0.236	0.058	0.259	0.048	0.289
MX18	0.202	0.105	0.239	0.074	0.189	0.210
MX19	0.217	0.142	0.248	0.077	0.288	0.028
MX20	0.271	0.018	0.280	0.011	0.232	0.188
MX21	0.208	0.096	0.235	0.081	0.174	0.206
MX22	0.069	0.274	0.108	0.240	0.111	0.197
MX23	0.081	0.283	0.127	0.199	0.109	0.201
MX24	0.086	0.265	0.136	0.189	0.189	0.136
MX25	0.151	0.228	0.223	0.066	0.323	0.009
MX26	0.171	0.180	0.190	0.133	0.135	0.191
MX27	0.127	0.227	0.190	0.111	0.231	0.114
MX28	0.154	0.173	0.157	0.140	0.190	0.186
MX29	0.264	0.086	0.288	0.020	0.312	0.031
MX30	0.242	0.135	0.246	0.043	0.282	0.053
MX31	0.239	0.062	0.258	0.061	0.215	0.165
MX32	0.128	0.207	0.194	0.078	0.284	0.108
MX33	0.200	0.193	0.210	0.118	0.246	0.032
MX34	0.187	0.176	0.188	0.157	0.129	0.163
MX35	0.263	0.029	0.289	0.000	0.276	0.142
MX36	0.163	0.248	0.158	0.165	0.230	0.035
MX37	0.275	0.135	0.238	0.072	0.279	0.000
MX38	0.162	0.220	0.193	0.130	0.263	0.032
MX39	0.232	0.100	0.247	0.055	0.189	0.177
MX40	0.204	0.186	0.210	0.089	0.223	0.088
MX41	0.112	0.252	0.119	0.252	0.038	0.227
MX42	0.055	0.307	0.144	0.113	0.278	0.103

APPENDIX J

DATA SET H

Data set H is a subset of data set D. H contains four of the texture variables from D that were shown in Section IV to mix linearly, even when the mean intensities of image halves were not equal, plus the variable MEAN. The five variables in Table J-1 (HOM3, HOM15, LM2, LM3, and MEAN) have been SIFTed to row sums of 1.0.

TABLE J-1. DATA SET H

	HOM3	HOM15	LM2	LM3	MEAN
MX01	0.039	0.381	0.000	0.406	0.175
MX02	0.000	0.422	0.008	0.415	0.155
MX03	0.011	0.448	0.053	0.345	0.143
MX04	0.219	0.299	0.263	0.198	0.021
MX05	0.292	0.223	0.318	0.146	0.020
MX06	0.317	0.203	0.347	0.089	0.044
MX07	0.307	0.036	0.321	0.011	0.325
MX08	0.289	0.034	0.330	0.011	0.335
MX09	0.327	0.000	0.327	0.023	0.324
MX10	0.225	0.195	0.283	0.064	0.233
MX11	0.221	0.189	0.247	0.112	0.230
MX12	0.312	0.141	0.327	0.087	0.133
MX13	0.216	0.258	0.230	0.163	0.134
MX14	0.288	0.235	0.222	0.102	0.153
MX15	0.179	0.275	0.213	0.180	0.153
MX16	0.218	0.181	0.256	0.072	0.273
MX17	0.148	0.319	0.078	0.351	0.103
MX18	0.257	0.133	0.303	0.094	0.212
MX19	0.279	0.182	0.318	0.099	0.123
MX20	0.345	0.023	0.356	0.014	0.262
MX21	0.243	0.112	0.275	0.095	0.274
MX22	0.079	0.312	0.123	0.273	0.213
MX23	0.109	0.378	0.170	0.267	0.076
MX24	0.113	0.348	0.179	0.248	0.112
MX25	0.179	0.272	0.265	0.079	0.205
MX26	0.213	0.224	0.236	0.166	0.160
MX27	0.155	0.278	0.232	0.136	0.199
MX28	0.191	0.214	0.194	0.173	0.228
MX29	0.307	0.099	0.334	0.023	0.236
MX30	0.306	0.171	0.311	0.055	0.158
MX31	0.289	0.074	0.311	0.074	0.252
MX32	0.156	0.252	0.237	0.095	0.260
MX33	0.265	0.255	0.278	0.156	0.046
MX34	0.248	0.234	0.249	0.208	0.062
MX35	0.303	0.034	0.333	0.000	0.350
MX36	0.219	0.334	0.212	0.222	0.013
MX37	0.378	0.186	0.327	0.099	0.009
MX38	0.230	0.312	0.273	0.184	0.000
MX39	0.333	0.143	0.354	0.079	0.091
MX40	0.283	0.259	0.292	0.124	0.042
MX41	0.142	0.320	0.152	0.320	0.066
MX42	0.070	0.391	0.184	0.144	0.212

APPENDIX K

DATA SET I

Data set I consists of 42 synthetic mixture images labeled NX1 through NX42. These mixtures correspond to samples MX1 through MX42 in data set D (Appendix F), except that mean intensity levels in the components of the mixture are not adjusted to a given level prior to mosaicking, as was the case for data set D. Table K-1 contains numerical values for each sample in I for mean intensity and for each of the six texture measures, which were shown in Section IV to mix linearly when mean intensities of mixture components vary (HOM3, HOM15, LM2, LM3, LM4, and LM5). Formatting and scaling of variables in Table K-1 is the same as in Table C-1.

Table K-1 also contains the compositions of each mixture by ice type. These are given at the bottom of the table.

TABLE K-1. VALUES OF TEXTURE MEASURES FOR DATA SET I

MEAS.	S.F.	NX1	NX2	NX3	NX4	NX5	NX6	NX7	NX8
MEAN	10^2	1.200	1.286	1.355	1.728	1.736	1.656	0.418	0.413
HOM3	10^{-1}	1.273	1.226	1.237	1.456	1.508	1.555	1.658	1.629
HOM15	10^{-1}	0.741	0.684	0.676	0.959	1.057	1.052	1.345	1.381
LM2	10^{16}	0.883	0.832	0.841	0.407	0.347	0.259	0.106	0.104
LM3	10^{13}	1.331	1.265	1.146	0.793	0.647	0.574	0.401	0.387
LM4	10^{12}	1.056	0.991	0.954	0.711	0.698	0.708	0.687	0.677
LM5	10^9	2.884	2.709	2.830	2.134	2.234	2.543	2.625	2.660
% FY		0	0	0	0	0	0	100	100
% SY		0	0	0	100	100	100	0	0
% MY		100	100	100	0	0	0	0	0

TABLE K-1. continued

MEAS.	S.F.	NX9	NX10	NX11	NX12	NX13	NX14	NX15	NX16
MEAN	10^2	0.427	0.931	0.939	1.332	1.323	1.190	1.277	0.783
HOM3	10^{-1}	1.692	1.501	1.496	1.590	1.480	1.451	1.419	1.495
HOM15	10^{-1}	1.443	1.003	1.024	1.116	0.927	0.897	0.876	1.056
LM2	10^{16}	0.095	0.316	0.362	0.422	0.577	0.543	0.892	0.331
LM3	10^{13}	0.434	0.550	0.651	0.741	0.870	0.758	1.078	0.548
LM4	10^{13}	0.740	0.655	0.742	0.821	0.848	0.792	0.898	0.654
LM5	10^9	2.742	2.212	2.430	2.598	2.567	2.514	2.720	2.276
% FY		100	50	50	25	25	25	25	50
% SY		0	25	25	50	50	25	25	0
% MY		0	25	25	25	25	50	50	50

TABLE K-1. continued

MEAS.	S.F.	NX17	NX18	NX19	NX20	NX21	NX22	NX23	NX24
MEAN	10 ²	1.446	1.010	1.387	0.785	0.728	1.029	1.556	1.445
HOM3	10 ⁻¹	1.393	1.534	1.530	1.668	1.541	1.310	1.337	1.341
HOM15	10 ⁻¹	0.842	1.112	1.056	1.348	1.183	0.795	0.774	0.824
LM2	10 ¹⁶	0.806	0.243	0.426	0.354	0.386	0.614	0.557	0.538
LM3	10 ¹³	1.271	0.609	0.691	0.579	0.711	1.030	0.917	0.904
LM4	10 ¹³	1.037	0.790	0.716	0.807	0.861	0.892	0.882	0.812
LM5	10 ⁹	2.947	2.646	2.304	2.925	2.794	2.539	2.539	2.457
% FY		0	50	25	75	75	25	0	0
% SY		50	50	75	25	0	0	75	25
% MY		50	0	0	0	25	75	25	75

TABLE K-1. continued

MEAS.	S.F.	NX25	NX26	NX27	NX28	NX29	NX30	NX31	NX32
MEAN	10 ²	1.077	1.214	1.105	1.004	0.894	1.239	0.830	0.883
HOM3	10 ⁻¹	1.429	1.482	1.409	1.435	1.618	1.584	1.589	1.407
HOM15	10 ⁻¹	0.879	0.957	0.846	0.955	1.214	1.091	1.252	0.936
LM2	10 ¹⁶	0.343	0.383	0.412	0.487	0.151	0.250	0.197	0.398
LM3	10 ¹⁸	0.556	0.776	0.689	0.771	0.427	0.499	0.559	0.594
LM4	10 ¹³	0.613	0.847	0.728	0.787	0.601	0.641	0.724	0.655
LM5	10 ⁹	2.118	2.572	2.364	2.566	2.172	2.227	2.516	2.351
% FY		33	33	33	33	67	33	67	33
% SY		33	33	32	33	33	67	0	0
% MY		33	33	33	33	0	0	33	67

TABLE K-1. continued

MEAS.	S.F.	NX33	NX34	NX35	NX36	NX37	NX38	NX39	NX40
MEAN	10^2	1.646	1.579	0.448	1.753	1.765	1.795	1.492	1.660
HOM3	10^{-1}	1.510	1.521	1.651	1.468	1.640	1.460	1.608	1.540
HOM15	10^{-1}	0.973	1.019	1.417	0.911	1.167	0.989	1.230	1.034
LM2	10^{16}	0.360	0.377	0.103	0.485	0.280	0.414	0.204	0.336
LM3	10^{18}	0.709	0.859	0.367	0.845	0.581	0.734	0.547	0.638
LM4	10^{13}	0.707	0.851	0.626	0.729	0.652	0.693	0.754	0.723
LM5	10^9	2.174	2.511	2.466	2.180	2.097	2.171	2.564	2.315
% FY		0	0	100	0	0	0	0	0
% SY		67	33	0	100	100	100	100	100
% MY		33	67	0	0	0	0	0	0

TABLE K-1. continued

MEAS.	S.F.	NX41	NX42
MEAN	10^2	1.561	1.104
HOM3	10^{-1}	1.405	1.314
HOM15	10^{-1}	0.863	0.791
LM2	10^{16}	0.557	0.525
LM3	10^{13}	1.143	0.686
LM4	10^{13}	0.986	0.674
LM5	10^9	2.665	2.333
% FY		0	0
% SY		9	0
% MY		100	100

APPENDIX L

DATA SET J

Data set J is a subset of data set I. J contains only the texture variables from I that were shown in Section IV to mix linearly, even when the mean intensities of the two image halves were not equal. The six texture measures in Table L-1 (HOM3, HOM15, LM2, LM3, LM4, and LM5) have been SIFTed to row sums of 1.0.

TABLE L-1. DATA SET J

	HOM3	HOM15	LM2	LM3	LM4	LM5
NX01	0.034	0.310	0.004	0.339	0.000	0.514
NX02	0.000	0.346	0.026	0.326	0.050	0.252
NX03	0.008	0.335	0.021	0.271	0.075	0.289
NX04	0.166	0.212	0.204	0.148	0.255	0.015
NX05	0.200	0.166	0.226	0.096	0.260	0.053
NX06	0.201	0.145	0.226	0.061	0.218	0.149
NX07	0.264	0.036	0.281	0.010	0.231	0.177
NX08	0.251	0.023	0.287	0.006	0.241	0.192
NX09	0.284	0.000	0.284	0.020	0.197	0.215
NX10	0.191	0.165	0.234	0.061	0.285	0.044
NX11	0.183	0.172	0.210	0.093	0.218	0.124
NX12	0.237	0.130	0.179	0.118	0.157	0.179
NX13	0.173	0.214	0.126	0.166	0.145	0.176
NX14	0.155	0.229	0.141	0.130	0.187	0.158
NX15	0.139	0.249	0.000	0.248	0.117	0.247
NX16	0.188	0.164	0.229	0.061	0.288	0.069
NX17	0.111	0.243	0.033	0.290	0.013	0.310
NX18	0.195	0.127	0.240	0.074	0.173	0.191
NX19	0.213	0.164	0.191	0.110	0.244	0.079
NX20	0.272	0.036	0.193	0.063	0.157	0.279
NX21	0.208	0.104	0.195	0.110	0.132	0.252
NX22	0.061	0.287	0.119	0.234	0.123	0.177
NX23	0.079	0.290	0.140	0.190	0.127	0.173
NX24	0.082	0.268	0.147	0.185	0.178	0.140
NX25	0.143	0.241	0.226	0.064	0.319	0.008
NX26	0.168	0.194	0.196	0.130	0.141	0.171
NX27	0.120	0.249	0.193	0.107	0.231	0.101
NX28	0.142	0.202	0.151	0.133	0.187	0.175
NX29	0.261	0.093	0.289	0.019	0.311	0.027
NX30	0.238	0.142	0.249	0.042	0.282	0.047
NX31	0.235	0.075	0.263	0.060	0.220	0.148
NX32	0.126	0.214	0.201	0.076	0.286	0.097
NX33	0.196	0.198	0.215	0.114	0.247	0.029
NX34	0.193	0.169	0.197	0.156	0.137	0.146
NX35	0.275	0.010	0.299	0.000	0.285	0.131
NX36	0.171	0.228	0.168	0.163	0.237	0.032
NX37	0.284	0.115	0.246	0.071	0.284	0.000
NX38	0.170	0.200	0.203	0.129	0.270	0.029
NX39	0.244	0.083	0.257	0.056	0.198	0.163
NX40	0.212	0.168	0.220	0.089	0.231	0.081
NX41	0.121	0.237	0.132	0.253	0.048	0.210
NX42	0.064	0.288	0.156	0.112	0.285	0.094

APPENDIX M

DATA SET K

Data set K is a subset of data set I. K contains four of the texture variables from I that were shown in Section IV to mix linearly, even when the mean intensities of image halves were not equal, plus the variable MEAN. The five variables in Table M-1 (HOM3, HOM15, LM2, LM3, and MEAN) have been SIFTed to row sums of 1.0.

TABLE M-1. DATA SET K

	HOM3	HOM15	LM2	LM3	MEAN
NX01	0.085	0.336	0.004	0.367	0.209
NX02	0.000	0.377	0.029	0.354	0.240
NX03	0.021	0.383	0.025	0.310	0.261
NX04	0.301	0.168	0.162	0.117	0.253
NX05	0.363	0.132	0.179	0.076	0.250
NX06	0.401	0.126	0.197	0.053	0.223
NX07	0.649	0.039	0.301	0.011	0.001
NX08	0.645	0.026	0.321	0.007	0.000
NX09	0.680	0.000	0.296	0.021	0.003
NX10	0.421	0.178	0.225	0.059	0.117
NX11	0.414	0.170	0.207	0.092	0.118
NX12	0.464	0.110	0.153	0.100	0.172
NX13	0.358	0.192	0.113	0.149	0.188
NX14	0.344	0.221	0.136	0.126	0.174
NX15	0.311	0.242	0.000	0.242	0.205
NX16	0.443	0.169	0.235	0.063	0.090
NX17	0.242	0.231	0.032	0.276	0.220
NX18	0.440	0.125	0.236	0.073	0.125
NX19	0.413	0.139	0.161	0.093	0.194
NX20	0.628	0.036	0.195	0.063	0.078
NX21	0.499	0.109	0.204	0.115	0.073
NX22	0.151	0.308	0.127	0.251	0.163
NX23	0.169	0.269	0.130	0.176	0.256
NX24	0.181	0.259	0.142	0.178	0.239
NX25	0.323	0.237	0.222	0.063	0.155
NX26	0.357	0.179	0.181	0.120	0.164
NX27	0.279	0.253	0.196	0.109	0.163
NX28	0.341	0.211	0.168	0.139	0.142
NX29	0.541	0.084	0.260	0.017	0.098
NX30	0.469	0.122	0.214	0.056	0.159
NX31	0.524	0.073	0.256	0.058	0.088
NX32	0.324	0.241	0.226	0.086	0.124
NX33	0.356	0.156	0.170	0.090	0.227
NX34	0.363	0.138	0.161	0.127	0.211
NX35	0.666	0.011	0.315	0.000	0.008
NX36	0.309	0.180	0.132	0.128	0.251
NX37	0.467	0.082	0.176	0.051	0.224
NX38	0.309	0.159	0.161	0.102	0.268
NX39	0.472	0.070	0.216	0.047	0.196
NX40	0.390	0.135	0.176	0.071	0.228
NX41	0.239	0.205	0.114	0.218	0.225
NX42	0.168	0.330	0.179	0.129	0.194

APPENDIX N

DATA SET L

Data set L is a SIFTed version of data set I.

TABLE N-1. DATA SET L

	HOM3	HOM15	LM2	LM3	LM4	LM5	MEAN
NX01	0.030	0.270	0.003	0.296	0.000	0.274	0.127
NX02	0.000	0.307	0.023	0.289	0.044	0.223	0.114
NX03	0.007	0.303	0.019	0.245	0.068	0.261	0.096
NX04	0.163	0.209	0.201	0.146	0.251	0.014	0.016
NX05	0.197	0.164	0.222	0.095	0.256	0.052	0.014
NX06	0.195	0.141	0.220	0.059	0.212	0.145	0.028
NX07	0.206	0.028	0.219	0.008	0.180	0.138	0.221
NX08	0.194	0.018	0.222	0.005	0.187	0.149	0.225
NX09	0.222	0.000	0.222	0.015	0.154	0.168	0.219
NX10	0.159	0.154	0.194	0.051	0.237	0.036	0.168
NX11	0.153	0.144	0.176	0.078	0.182	0.103	0.164
NX12	0.215	0.118	0.163	0.107	0.142	0.163	0.092
NX13	0.156	0.193	0.113	0.150	0.131	0.159	0.098
NX14	0.136	0.201	0.123	0.114	0.164	0.138	0.123
NX15	0.124	0.221	0.000	0.220	0.104	0.219	0.112
NX16	0.152	0.133	0.185	0.049	0.233	0.055	0.193
NX17	0.103	0.225	0.031	0.269	0.012	0.287	0.073
NX18	0.167	0.109	0.206	0.063	0.148	0.163	0.144
NX19	0.194	0.150	0.174	0.100	0.222	0.072	0.088
NX20	0.225	0.029	0.160	0.052	0.130	0.231	0.173
NX21	0.168	0.084	0.158	0.089	0.106	0.204	0.192
NX22	0.052	0.242	0.100	0.197	0.103	0.149	0.159
NX23	0.075	0.275	0.132	0.180	0.120	0.164	0.054
NX24	0.076	0.247	0.136	0.170	0.164	0.130	0.077
NX25	0.122	0.206	0.193	0.055	0.272	0.007	0.145
NX26	0.149	0.172	0.173	0.115	0.125	0.152	0.114
NX27	0.103	0.215	0.166	0.092	0.199	0.087	0.138
NX28	0.120	0.171	0.136	0.112	0.159	0.148	0.154
NX29	0.217	0.077	0.240	0.016	0.258	0.023	0.168
NX30	0.211	0.126	0.221	0.038	0.251	0.042	0.111
NX31	0.194	0.062	0.217	0.050	0.182	0.123	0.174
NX32	0.104	0.177	0.166	0.063	0.235	0.080	0.176
NX33	0.190	0.191	0.208	0.111	0.239	0.028	0.034
NX34	0.184	0.161	0.188	0.149	0.131	0.142	0.045
NX35	0.213	0.008	0.231	0.000	0.220	0.101	0.227
NX36	0.169	0.226	0.167	0.162	0.234	0.032	0.010
NX37	0.282	0.114	0.244	0.071	0.282	0.000	0.007
NX38	0.170	0.200	0.203	0.129	0.270	0.029	0.000
NX39	0.229	0.078	0.241	0.052	0.185	0.153	0.061
NX40	0.206	0.163	0.213	0.086	0.224	0.078	0.030
NX41	0.114	0.226	0.125	0.240	0.046	0.199	0.050
NX42	0.055	0.247	0.134	0.096	0.244	0.081	0.145

APPENDIX O

DATA SET M

Data set M consists of 10 mixture images. Samples BX1 through BX4 are images of natural mixtures of first-year and multiyear ice. Samples BX5 through BX10 are large synthetic mixtures constructed by mosaicking samples from data set I. Table O-1 contains numerical values for each sample in M for mean intensity and for each of the six texture measures, which were shown in Section IV to mix linearly when mean intensities of mixture components vary (HOM3, HOM15, LM2, LM3, LM4, LM5). Formatting and scaling of variables in Table O-1 is the same as in Table C-1.

Table K-1 also contains compositions of each mixture by ice type. These are given at the bottom of the table.

TABLE O-1. VALUES OF TEXTURE MEASURES FOR DATA SET M

MEAS.	S.F.	BX1	BX2	BX3	BX4	BX5	BX6	BX7	BX8
MEAN	10^2	0.508	0.999	0.720	0.839	0.427	1.721	1.284	1.073
HOM3	10^{-1}	1.564	1.295	1.458	1.319	1.653	1.484	1.277	1.564
HOM15	10^{-1}	1.262	0.781	1.100	0.867	1.384	0.956	0.731	1.152
LM2	10^{16}	0.375	1.016	0.495	0.945	0.104	0.384	0.783	0.245
LM3	10^{13}	0.623	1.138	0.742	0.992	0.410	0.724	1.161	0.576
LM4	10^{13}	0.738	0.912	0.779	0.839	0.692	0.735	0.950	0.723
LM5	10^9	2.578	2.641	2.591	2.563	2.627	2.321	2.685	2.493
% FY		90	40	69	60	100	0	0	50
% SY		0	0	0	0	0	100	0	50
% MY		10	60	31	40	0	0	100	0

TABLE O-1. continued

MEAS.	S.F.	BX9	BX10
MEAN	10^2	0.889	1.525
HOM3	10^{-1}	1.462	1.378
HOM15	10^{-1}	1.044	0.838
LM2	10^{16}	0.441	0.585
LM3	10^{13}	0.808	0.976
LM4	10^{13}	0.839	0.868
LM5	10^9	2.681	2.544
% FY		50	0
% SY		0	50
% MY		50	50

APPENDIX P

DATA SET N

Data set N is a subset of data set M. N contains MEAN plus four of the texture variables from N, which were shown in Section IV to mix linearly, even when the mean intensities of the two image halves were not equal. The five texture measures in Table P-1 (HOM3, HOM15, LM2, LM3, and MEAN) have been SIFTed to row sums of 1.0.

TABLE P-1. DATA SET N

	HOM3	HOM15	LM2	LM3	MEAN
BX01	0.266	0.065	0.245	0.099	0.326
BX02	0.019	0.370	0.000	0.388	0.223
BX03	0.178	0.161	0.211	0.164	0.286
BX04	0.046	0.325	0.032	0.318	0.280
BX05	0.333	0.000	0.333	0.000	0.333
BX06	0.238	0.283	0.299	0.180	0.000
BX07	0.000	0.386	0.099	0.386	0.130
BX08	0.284	0.132	0.315	0.082	0.186
BX09	0.175	0.185	0.224	0.188	0.228
BX10	0.108	0.337	0.190	0.304	0.061

APPENDIX Q

DATA SET O

Data set O is a subset of data set M. O contains only the texture variables from M, which were shown in Section IV to mix linearly, even when the mean intensities of the two image halves were not equal. The six texture measures in Table Q-1 (HOM3, HOM15, LM2, LM3, LM4, and LM5) have been SIFTed to row sums of 1.0.

TABLE Q-1. DATA SET O

	HOM3	HOM15	LM2	LM3	LM4	LM5
BX01	0.220	0.054	0.203	0.082	0.237	0.204
BX02	0.016	0.311	0.000	0.327	0.050	0.296
BX03	0.144	0.130	0.171	0.133	0.199	0.222
BX04	0.039	0.278	0.027	0.272	0.151	0.233
BX05	0.260	0.000	0.260	0.000	0.260	0.219
BX06	0.175	0.208	0.220	0.133	0.265	0.000
BX07	0.000	0.307	0.078	0.307	0.000	0.307
BX08	0.216	0.100	0.239	0.062	0.249	0.134
BX09	0.137	0.145	0.176	0.148	0.120	0.275
BX10	0.082	0.256	0.145	0.231	0.097	0.188

Distribution List

Asst Secretary of the Navy
(Research, Engineering & Systems)
Navy Department
Washington DC 20350-1000

Chief of Naval Operations
Navy Department
Washington DC 20350-1000
Attn: OP-02
OP-71
OP-0962X
OP-987

Oceanographer of the Navy
Chief of Naval Operations
U.S. Naval Observatory
341h & Mass Ave., NW
Washington DC 20390-1800
Attn: OP-96

Commander
Naval Air Development Center
Warminster PA 18974-5000

Commanding Officer
Naval Coastal Systems Center
Panama City FL 32407-5000

Commander
Space & Naval Warfare Sys Com
Washington DC 20363-5100

Commanding Officer
Naval Environmental Prediction
Research Facility
Monterey CA 93943-5006

Commander
Naval Facilities Eng Command
Naval Facilities Eng Command HQ
200 Stovall St.
Alexandria VA 22332-2300

Commanding Officer
Naval Ocean R&D Activity
Stennis Space Center MS 39529-5004
Attn: Code 100
Code 105
Code 115
Code 117, J. Hammack
Code 125EX
Code 125L (13)
Code 125P (1)
Code 200
Code 300
Code 350, D. Hickman
Code 351, J. Byrnes
Code 351, J. Braud
Code 351, M. Lohrenz

Brooke Farquhar
NORDA Liaison Office
Crystal Plaza #5, Room 802
2211 Jefferson Davis Hwy.
Arlington VA 22202-5000

Commanding Officer
Naval Research Laboratory
Washington DC 20375

Commander
Naval Oceanography Command
Stennis Space Center MS 39529-5000

Commanding Officer
Fleet Numerical Oceanography Center
Monterey CA 93943-5005

Commanding Officer
Naval Oceanographic Office
Stennis Space Center MS 39522-5001
Attn: Code GGAP, B. Muller
Code A, J. Depner

Commander
Naval Ocean Systems Center
San Diego CA 92152-5000

Commanding Officer
ONR Branch Office
Box 39
FPO New York NY 09510-0700

Commander
David W. Taylor Naval Research Center
Bethesda MD 20084-5000

Commander
Naval Surface Weapons Center
Dahlgren VA 22448-5000

Commanding Officer
Naval Underwater Systems Center
Newport RI 02841-5047

Superintendent
Naval Postgraduate School
Monterey CA 93943

Director of Navy Laboratories
Rm 1062, Crystal Plaza Bldg 5
Department of the Navy
Washington DC 20360

Officer in Charge
New London Laboratory
Naval Underwater Sys Cen Det
New London CT 06320

Director
National Ocean Data Center
WSC1 Room 103
6001 Executive Blvd.
Rockville MD 20852
Attn: G. W. Withee

Director
Woods Hole Oceanographic Inst
P.O. Box 32
Woods Hole MA 02543

University of California
Scripps Institute of Oceanography
P.O. Box 6049
San Diego CA 92106

Officer in Charge
Naval Surface Weapons Center Det
White Oak Laboratory
10901 New Hampshire Ave.
Silver Spring MD 20903-5000
Attn: Library

Commanding Officer
Fleet Anti-Sub Warfare Training Center,
Atlantic
Naval Station
Norfolk VA 23511-6495

Defense Mapping Agency Sys Cen
12100 Sunset Hill Rd. #200
Reston VA 22090-3207
Attn: SGWN
Mel Wagner
Ed Danford

Office of Naval Technology
800 N. Quincy St.
Arlington VA 22217-5000
Attn: Code 20, Dr. P. Selwyn
Code 228, Dr. M. Briscoe
Code 234, Dr. C. V. Votaw

Office of Naval Research
800 N. Quincy St.
Arlington VA 22217-5000
Attn: Code 10
Code 10D/10P, Dr. E. Silva
Code 12
Code 112, Dr. E. Hartwig

Commander
Naval Sea Systems Command
Naval Sea Systems Command HQ
Washington DC 20362-5101

Commanding Officer
Naval Civil Engineering Laboratory
Port Hueneme CA 93043

Commander
Naval Air Systems Command
Naval Air Systems Command HQ
Washington DC 20361-0001

Pennsylvania State University
Applied Research Laboratory
P.O. Box 30
State College PA 16801

University of Texas at Austin
Applied Research Laboratories
P.O. Box 8029
Austin TX 78713-6029

Johns Hopkins University
Applied Physics Laboratory
Johns Hopkins Rd.
Laurel MD 20707

University of Washington
Applied Physics Laboratory
1013 Northeast 40th St.
Seattle WA 98105

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. Agency Use Only (Leave blank).		2. Report Date. March 1990	3. Report Type and Dates Covered. Final	
4. Title and Subtitle. A Global Approach to Image Texture Analysis			5. Funding Numbers. Program Element No. 62435N Project No. Task No. Accession No.	
6. Author(s). R. J. Holyer				
7. Performing Organization Name(s) and Address(es). Ocean Science Directorate Naval Ocean Research and Development Activity Stennis Space Center, Mississippi 39529-5004			8. Performing Organization Report Number. NORDA Report 238	
9. Sponsoring/Monitoring Agency Name(s) and Address(es).			10. Sponsoring/Monitoring Agency Report Number.	
11. Supplementary Notes.				
12a. Distribution/Availability Statement. Approved for public release; distribution is unlimited. Naval Ocean Research and Development Activity, Stennis Space Center, Mississippi 39529-5004.			12b. Distribution Code.	
13. Abstract (Maximum 200 words). A new approach to image texture analysis is developed. The approach is based on linear unmixing of texture measures calculated over an entire image (called a global approach), as opposed to most present texture analysis techniques that compute texture over small neighborhoods (called a local approach). The new global paradigm is appropriate for images where spatial scales of the texture variability are large with respect to the pixel spacing, thereby making the local approach ineffective. Airborne passive microwave imagery of Arctic sea ice contain textures that vary with ice-type. These ice textures are of the type best treated by the global approach. Sea-ice imagery are used as test data to evaluate the global techniques that are developed. Pure, single ice-type images; synthetic mixtures formed by mosaicking pure ice-type subimages in known proportions; and naturally occurring mixture images are analyzed in the course of the study. Proportions of first-year, second-year, and multiyear ice within mixture images are retrieved with root-mean-square accuracies as low as 0.04 by the new method. This accuracy is adequate to be useful in many Arctic studies, but more important, the global technique seems promising for many other remote sensing and general image processing applications. Research areas that are required to advance the global method are enumerated. The most important advancement in support of the global method would be the development of new image texture measures that exhibit linear properties under mixing operations.				
14. Subject Terms. remote sensing, satellite image interpretation, image texture analysis			15. Number of Pages. 286	
			16. Price Code.	
17. Security Classification of Report. Unclassified	18. Security Classification of This Page. Unclassified	19. Security Classification of Abstract. Unclassified	20. Limitation of Abstract. None	